Connecting Satellite Observations with Water Cycle Variables through Land Data Assimilation: Examples Using the NASA GEOS-5 LDAS

Rolf H. Reichle[1], Gabriëlle J. M. De Lannoy[1,2], Barton A. Forman[3], Clara S. Draper[1,2], and Qing Liu[1,4]

[1]*Global Modeling and Assimilation Office, NASA/GSFC, Greenbelt, MD, USA.*

[2]*Universities Space Research Association, Columbia, MD, USA.*

[3]*Dept. of Civil and Environmental Engineering, University of Maryland, College Park, MD, USA.*

[4]*Science Systems and Applications, Inc., Lanham, MD, USA.*

Abstract A land data assimilation system (LDAS) can merge satellite observations (or retrievals) of land surface hydrological conditions, including soil moisture, snow, and terrestrial water storage (TWS), into a numerical model of land surface processes. In theory, the output from such a system is superior to estimates based on the observations or the model alone, thereby enhancing our ability to understand, monitor, and predict key elements of the terrestrial water cycle. In practice, however, satellite observations do not correspond directly to the water cycle variables of interest. The present paper addresses various aspects of this seeming mismatch using examples drawn from recent research with the ensemble-based NASA GEOS-5 LDAS. These aspects include (i) the assimilation of coarse-scale observations into higher-resolution land surface models, (ii) the partitioning of satellite observations (such as TWS retrievals) into their constituent water cycle components, (iii) the forward modeling of microwave brightness temperatures over land for radiance-based soil moisture and snow assimilation, and (iv) the selection of the most relevant types of observations for the analysis of a specific water cycle variable that is not observed (such as root zone soil moisture). The solution to these challenges involves the careful construction of an observation operator that maps from the land surface model variables of interest to the space of the assimilated observations.

Keywords *Land data assimilation, land surface modeling, satellite remote sensing, soil moisture, snow, terrestrial water storage, ensemble Kalman filter*

Corresponding Author: R. H. Reichle
Global Modeling and Assimilation Office (Code 610.1)
NASA Goddard Space Flight Center
8800 Greenbelt Road
Greenbelt, MD 20771, USA
Email: Rolf.Reichle@nasa.gov
Phone: +1-301-614-5693

1 Introduction

The water cycle plays a crucial role in Earth's climate and environment, yet there are still large gaps in our understanding of its components, particularly at the land surface (Bengtsson, 2013; Lahoz and De Lannoy, 2013; Trenberth 2013). Over the past decade there has been a steady increase in the number and types of satellite observations (or retrievals) related to land surface hydrological conditions, including soil moisture, snow, and terrestrial water storage (TWS; Bartalis et al. 2007; Bruinsma et al. 2010; Clifford 2010; de Jeu et al. 2008; Entekhabi et al. 2010; Foster et al. 2005, 2011; Gao et al. 2010; Hall and Riggs 2007; Hall et al. 2010; Horwath et al. 2011; Kelly 2009; Kerr et al. 2010; Li et al. 2007; Liu et al. 2011b; Njoku et al. 2003; Parinussa et al. 2012; Peltier 2013;Pulliainen 2006; Rowlands et al. 2005, 2010; Swenson and Wahr 2006; Tedesco and Narvekar 2010; Tedesco et al. 2010; Wahr et al. 2004).

These observations can be assimilated into land surface models to provide land surface hydrological estimates that are generally superior to the satellite observations or model estimates alone (Andreadis and Lettenmaier 2006; Crow and Wood 2003; De Lannoy et al. 2012a; de Rosnay et al. 2012a, 2012b; Draper et al. 2012; Drusch 2007; Dunne and Entekhabi 2006; Durand and Margulis 2008; Forman et al. 2012a; Houborg et al. 2012; Li et al. 2012; Liu et al. 2011a; Margulis et al. 2002; Pan and Wood 2006; Pan et al. 2008; Reichle and Koster 2005; Reichle et al. 2007, 2009; Sahoo et al. 2012; Su et al. 2008, 2010; Zaitchik et al. 2008).

However, land data assimilation systems must be designed carefully such that a number of conceptual problems can be overcome and the potential improvements from data assimilation can be realized. Earlier work addressed the bias between the satellite observations and model estimates within the assimilation system (De Lannoy et al. 2007; Drusch et al. 2005; Kumar et al. 2012; Reichle and Koster 2004). Moreover, approaches to efficient error modeling within the assimilation system, including adaptive methods, needed to be developed (Crow and Reichle 2008; Crow and van den Berg 2010; Reichle et al. 2008b). An overview of some

relevant earlier literature in the context of the ensemble-based Goddard Earth Observing System Model, Version 5 (GEOS-5) land data assimilation system (LDAS) developed at the NASA Global Modeling and Assimilation Office (GMAO) is provided by Reichle et al. (2009).

Despite the early successes, the design and application of land data assimilation systems still face additional conceptual problems. While land surface models are flexible in the design and choice of model variables, satellite observations do not necessarily correspond directly to the water cycle variables of interest. For example, space-borne microwave observations can be converted into estimates of snow amount or surface soil moisture, but the spatial resolution of such microwave-based retrievals is usually much coarser than desired. Moreover, satellites typically observe electromagnetic properties such as backscatter and/or radiances (or brightness temperatures) that are only indirectly related to snow amounts or soil moisture levels. Furthermore, satellite-observed backscatter and radiances are at best sensitive to moisture in the top few centimeters of the soil. Information on important water cycle components such as root zone soil moisture must therefore be gained through even more indirect pathways in the land data assimilation system.

The present paper addresses several major challenges that all relate to a seeming mismatch between the assimilated observations and the water cycle variables of interest. This mismatch can be overcome through the careful design of the land data assimilation system. The conceptual challenges discussed here can be summarized as follows:

(i) How can coarse-scale satellite observations increase our knowledge of land surface conditions at finer scales (horizontal downscaling), and how can unobserved areas be updated using information from neighboring observations?

(ii) How can vertically integrated measurements (such as TWS) be partitioned into their component variables within the assimilation system?

(iii) How can satellite radiances (rather than geophysical retrievals) be assimilated to improve estimates of land surface hydrological conditions (e.g., soil moisture and snow)?

(iv) How can the most relevant types of observations be selected for the analysis of a water cycle component that is not observed (such as root zone soil moisture)?

The present paper illustrates each of these conceptual problems based on recent progress using the GEOS-5 system for land surface hydrological data assimilation. The examples use satellite observations of land surface water cycle components from the Advanced Microwave Scanning Radiometer for EOS (AMSR-E), the Moderate Resolution Imaging Spectroradiometer (MODIS), the Gravity Recovery and Climate Experiment (GRACE) mission, the Advanced Scatterometer (ASCAT), and the Soil Moisture Ocean Salinity (SMOS) mission for the analysis of soil moisture (AMSR-E, ASCAT, SMOS, GRACE), snow (AMSR-E, MODIS, GRACE), and TWS (GRACE). After a brief discussion of the GEOS-5 LDAS, section 2 provides details and references for the various satellite observations used in the examples. Section 3 addresses each of the above-mentioned challenge questions in a separate subsection. Results are discussed and summarized in section 4. Finally, section 5 provides conclusions and a brief outlook on future research directions.

2 Data and Methods

2.1 GEOS-5 Land Data Assimilation System

The GEOS-5 LDAS consists of the NASA Catchment land surface model and an implementation of the ensemble Kalman filter (EnKF; Evensen 2003). The GEOS-5 EnKF has also been included in the NASA Land Information System, a comprehensive land surface modeling and assimilation software framework, so that it can be used with a variety of land surface models (Kumar et al. 2008a, 2008b). A brief summary of the key characteristics of the system is provided below. For a more comprehensive discussion see Reichle et al. (2009) and references therein.

The Catchment land surface model (hereinafter Catchment model; Ducharne et al. 2000; Koster et al. 2000) differs from traditional, layer-based land surface models by including an explicit treatment of the spatial variation within each hydrological catchment (or computational element) of the soil water and water table depth, as well as its effect on runoff and evaporation. Within each element, the vertical profile of soil water down to the bedrock is given by the equilibrium soil moisture profile and the deviations from the equilibrium profile. The deviations are described by excess and deficit variables for a 0-2 cm (or 0-5 cm) surface layer and for a "root zone" layer that extends from the surface to a depth z_R of 75 cm $\leq z_R \leq$ 100 cm depending on local soil conditions. The spatial variability of soil moisture is diagnosed at each time step from the bulk water prognostic variables and the statistics of the catchment topography. One key feature of the Catchment model is the groundwater component implicit in the modeling of the water table depth (through the modeling of the subsurface water profile down to the bedrock). This groundwater component is critically important for the assimilation of TWS retrievals (section 3.2).

The Catchment model also includes a state-of-the-art, multi-layer, global snow model (Stieglitz et al. 2001). In each watershed, the evolution of the amount of water in the snow pack (or snow water equivalent; SWE), the snow depth, and the

snow heat content in response to surface meteorological conditions and snow compaction is modeled using three layers. The soil, vegetation, and snow model parameters used in the Catchment model are from the NASA GEOS-5 global modeling system (Rienecker et al. 2008).

The EnKF is a Monte-Carlo variant of the Kalman filter, which sequentially updates model forecasts in response to observations based on the relative uncertainty of the model and the observations. The key idea behind the EnKF is that the relevant parts of the model error covariance structure can be captured by a small ensemble of model trajectories. Each member of the ensemble experiences perturbed instances of the observed forcing fields (representing errors in the forcing data) and/or randomly generated noise that is added to the model parameters and prognostic variables (representing errors in model physics and parameters). The model error covariance matrices that are required for the filter update can then be diagnosed from the ensemble at the update time. The EnKF is flexible in its treatment of errors in model dynamics and parameters. It is also very suitable for modestly nonlinear problems and has become a popular choice for land data assimilation (Andreadis and Lettenmaier 2006; Durand and Margulis 2008; Kumar et al. 2008a, 2008b; Pan and Wood 2006; Reichle et al. 2002a, 2002b; Su et al. 2008; Zhou et al. 2006).

To realize the potential benefits from data assimilation, the assimilation system must be supplied with appropriate input parameters for the description of model and observation errors. For an ensemble-based system such as the GEOS-5 LDAS, for example, standard deviations, spatial and temporal correlations, and cross-correlations must be specified for the perturbations that are applied to each ensemble member. A detailed discussion of the error parameters in the examples discussed here is beyond the scope of the paper. The reader is referred to the references provided with each example as well as the overview discussion of Reichle et al. (2009).

2.2 Assimilated Observations

The data assimilation examples discussed in this paper use various types of satellite observations from a number of polar orbiting sensors/platforms, including passive and active microwave observations (AMSR-E, SMOS, and ASCAT), visible and near-infrared observations (MODIS), and gravimetric observations (GRACE).

AMSR-E, which operated with nominal performance between 2002 and 2011, is a scanning, dual polarization radiometer that measured microwave emission from the Earth at six frequencies (6.9, 10.7, 18.7, 23.9, 36.5, and 89.0 GHz), ranging in resolution from ~50 km at 6.9 GHz to ~5 km at 89.0 GHz (Knowles et al. 2006). Its successor, AMSR2, was launched in May 2012 (http://www.jaxa.jp/projects/sat/gcom_w/index_e.html). The training and validation of the empirical microwave radiative transfer model for snow-covered land surfaces in section 3.3.2 uses the 10.7, 18.7, and 36.5 GHz AMSR-E brightness temperatures, while the snow assimilation example in section 3.1 uses SWE retrievals that are based on the difference between the 18.7 GHz and the 36.5 GHz brightness temperatures (Kelly 2009). The soil moisture assimilation examples in section 3.4 use surface (top 1 cm) soil moisture retrievals that are derived from the 6.9 GHz and 10.7 GHz brightness temperatures (de Jeu et al. 2008; Njoku et al. 2003).

SMOS was launched in 2009 and its Microwave Imaging Radiometer with Aperture Synthesis (MIRAS) sensor provides multi-angular L-band (1.4 GHz) brightness temperature observations at horizontal and vertical polarization and a nominal spatial resolution of 43 km (Kerr et al. 2010). SMOS brightness temperatures are used in section 3.3.1.

ASCAT is a 5.3 GHz radar system that illuminates the Earth's surface and measures the energy scattered back to the instrument. The ASCAT surface (top 1 cm) soil moisture retrievals used in section 3.4.2 are derived from these backscatter measurements (Bartalis et al. 2007; Wagner et al. 1999) and are provided in units of degree of saturation.

MODIS (2000-present) provides visible and near-infrared observations from which snow cover fraction (SCF) can be retrieved under clear-sky conditions (Hall and Riggs 2007). High-resolution (500 m) MODIS SCF retrievals are in section 3.1.

Through the measurement of gravitational anomalies associated with the accumulation (or loss) of mass near the Earth's surface, GRACE provides approximately monthly, basin-scale (>150,000 km^2) estimates of variations in TWS, which includes snow, ice, surface water, soil moisture, and groundwater (Bruinsma et al. 2010; Horwath et al. 2011; Rodell et al. 2009; Rowlands et al. 2005, 2010; Swenson and Wahr 2006; Tang et al. 2010; Wahr et al. 2004). The assimilation experiments of section 3.2 use GRACE TWS retrievals.

2.3 Validation Data and Approach

For each of the examples presented in section 3, the output from the assimilation system was evaluated against independent data from various sources. In section 3.1, in situ SWE measurements from United States Department of Agriculture Snowpack Telemetry (SNOTEL; Schaefer et al. 2007) network sites in Colorado were used for evaluation, along with snow depth measurements from National Oceanic and Atmospheric Administration Cooperative Observer Program (COOP; http://www.ncdc.noaa.gov) sites.

SWE estimates for the Mackenzie River basin, used for evaluation in section 3.2, were derived from the daily snow depth product of the Canadian Meteorological Centre (CMC) daily snow analysis (Brasnett 1999; Brown and Brasnett 2010) at a horizontal resolution of approximately 24 km. The CMC snow analysis is based on optimal interpolation of in situ daily snow depth observations and aviation reports with a first-guess field generated from a snow model driven by output from the CMC weather model. Using the snow class map shown in Sturm et al. (1995), SWE estimates were obtained by multiplying the CMC snow depths with the Sturm et al. (2010) snow densities. Furthermore, runoff estimates for the

Mackenzie River basin and its major sub-basins provided by the Global Runoff Data Center (GRDC; http://www.bafg.de/GRDC) were used in section 3.2.

The radiative transfer models of section 3.3 were evaluated with AMSR-E and SMOS microwave brightness temperatures using a split sample approach in which one portion of the satellite brightness data was used for calibration or training and another, different portion was used for evaluation.

In situ profile soil moisture observations used for evaluation in section 3.4 are from the United States Department of Agriculture Soil Climate Analysis Network (SCAN) / SNOTEL (Schaefer et al. 2007) network in the contiguous US and from the Murrumbidgee Soil Moisture Monitoring Network (Smith et al. 2012) in Australia. Both sets of measurements were subjected to extensive quality control steps, including automatic detection of problematic observations and a visual inspection of the time series prior to using the data for evaluation.

Metrics used for skill assessment include the bias, root mean square error (RMSE), and time series correlation coefficient (R). When specified, anomalies were computed by removing a seasonally varying climatology from the data before computing the metrics.

3 Results

3.1 Assimilation of Sparse and Coarse-Scale Observations

Snow is an important component of the land system because of its strong impact on the land surface water and energy balance, weather, climate, and water resources (Barnett et al. 2005; Rott 2013). However, land surface models often represent snow processes poorly. Satellite observations of SWE can be retrieved from passive microwave sensors, but they are only available at relatively coarse resolution. Moreover, SWE retrievals, like most satellite observations, do not provide complete spatial and continuous temporal coverage due to orbit or sensor limitations. The challenge is therefore to design an assimilation system that can use coarse-scale satellite observations to provide enhanced model estimates at the finer scales of interest (horizontal downscaling) and that can also propagate the information to intermittently unobserved areas.

Using AMSR-E SWE retrievals and MODIS SCF observations, De Lannoy et al. (2010, 2012a) developed a data assimilation and downscaling technique for estimating fine-scale (1 km) snow fields using coarse-scale (25 km) SWE retrievals and fine-scale (500 m) SCF retrievals for a domain in Northern Colorado, USA. In their study, the authors used the LIS version of the GEOS-5 EnKF together with the Noah land surface model (Ek et al. 2003) (rather than the GEOS-5 LDAS and the Catchment model used elsewhere in this paper). The Noah model simulates a single snow layer with two prognostic variables for SWE and snow depth. The default LIS soil, vegetation, and general parameter tables for Noah were used, including a Noah-specific maximum snow albedo.

Figure 1 shows schematically how the coarse-scale SWE retrievals are used. The fine-scale model grid is represented by the dashed lines in the figure. The coarse-scale grid of the SWE observations is represented by the solid lines and light/dark gray shading, and the center points of individual SWE retrievals are marked with crosses. Let us now consider the analysis update of the fine-scale model grid cell indicated by the solid black square. First, it is important to emphasize that the

301 coarse-scale SWE retrievals are *not* compared directly to the SWE estimate at the
302 fine-scale model grid cell. Rather, the model SWE is aggregated to the coarse
303 grid of the retrievals, that is, the fine-scale model forecast is mapped into the
304 coarse-scale *observation space*. This aggregation is part of the *observation*
305 *operator* that maps the model states to the observations. Observation-minus-
306 model-forecast residuals (or *innovations*) are then computed at the coarse scale of
307 the observation space. The *Kalman gain* matrix transforms the (observation-
308 space) innovations into the (model-space) *increments*. It is computed from error
309 correlations between the model states at the fine scale and the model-predicted
310 measurements at the coarse scale. Finally, the increments are added to the (fine-
311 scale) model forecast in the *analysis update*. See De Lannoy et al. (2010) for a
312 discussion based on equations.
313
314 Second, multiple coarse-scale SWE retrievals in the vicinity of the fine-scale
315 model grid cell in question are used for the analysis update. Specifically, the
316 update uses the three coarse-scale SWE retrievals marked by black crosses that
317 are within a given radius (indicated by the white semi-circle) around the fine-scale
318 model grid cell in question (Figure 1). Note that this model grid cell would be
319 updated even if the SWE retrieval directly covering it were unavailable – the two
320 neighboring SWE retrievals (dark gray shading) would still contribute to the
321 update. The connection between the neighboring SWE retrievals and the model
322 grid cell in question relies on horizontal model *error* correlations that are due to,
323 for example, errors in large-scale model forcing fields such as snowfall or air
324 temperature.
325
326 To assimilate SCF, the Noah model snow depletion curve acts as the observation
327 operator that converts fine-scale modeled SWE into SCF estimates. Unlike binary
328 indicators of snow presence, the continuous SCF observations used here can thus
329 be assimilated with an EnKF, taking advantage of the distribution of SCF values
330 across the ensemble. Snow-free or fully snow covered conditions in the model
331 forecast ensemble were addressed by supplementing the EnKF with rule-based
332 update procedures (De Lannoy et al. 2012a). If at a given time and location all
333 members of the model forecast ensemble are snow-free but the SCF observation

indicates the presence of snow, then a nominal amount of snow is added to the model forecast. If all forecast ensemble members have full snow cover and the observed SCF indicates less than full cover, then the model forecast SWE and snow depth are reduced by a fixed fraction.

Figure 2 shows several observed and modeled snow fields for one snow season. The top row shows the coarse-scale (25 km) AMSR-E SWE retrievals, with data missing when the satellite swath does not fully cover the study area. MODIS fine-scale estimates of SCF, shown in the second row, are available only for clear-sky conditions. The bottom four rows of Figure 2 show that the assimilation of coarse-scale AMSR-E SWE and fine-scale MODIS SCF observations both result in realistic fine-scale spatial SWE patterns.

Through a quantitative validation of the assimilation results with independent measurements at individual SNOTEL and COOP sites over the course of 8 years, De Lannoy et al. (2012a) demonstrate improvements from the assimilation of SWE and/or SCF retrievals in shallow snow packs, but not in deep snow packs (not shown). The validation also shows that joint assimilation of SWE and SCF retrievals yields significantly improved RMSE and correlation values. For example, the RMSE for SWE versus COOP site measurements was reduced by 21% (from 78 mm to 62 mm) through the joint assimilation of satellite SWE and SCF retrievals. Furthermore, SCF assimilation was found to improve the timing of the onset of the snow season, albeit without a net improvement of SWE estimates. In areas of deep snow, however, AMSR-E retrievals are typically biased low and require bias correction (or scaling of the observations) prior to data assimilation. De Lannoy et al. (2012a) also showed that the interannual SWE variations could not be improved through the assimilation of AMSR-E because the AMSR-E retrievals lack realistic interannual variability in deep snow packs. These deficiencies in the AMSR-E SWE retrievals motivated the development of the empirical microwave radiative transfer model (section 3.3.2) towards a radiance-based snow analysis.

Of course, horizontal downscaling is not only important for snow assimilation. Low-frequency passive microwave brightness temperature observations such those from AMSR-E and SMOS (and the corresponding soil moisture retrievals) are at the coarse resolution of ~50 km. But for applications such as weather prediction, soil moisture estimates are needed at hydrometeorological scales of ~10 km or better. Examples of soil moisture downscaling based on data assimilation are provided by Reichle et al. (2001), Sahoo et al. (2012), and Zhou et al. (2006). Also, Reichle and Koster (2003) addressed the propagation of observational soil moisture information to unobserved regions.

3.2 Partitioning of Terrestrial Water Storage Observations

Passive microwave (e.g. AMSR-E) retrievals have been used in conjunction with land surface models to better characterize snow (section 3.1) and soil moisture (section 3.4). Gravimetric measurements such as from GRACE can provide monthly, basin-scale (>150,000 km^2) estimates of changes in TWS (section 2.2). Since TWS is vertically integrated and includes groundwater, soil moisture, snow, and surface water, TWS retrievals offer significant insights into the regional- and continental-scale water balance and, through data assimilation, the potential to learn more about hydrologic processes (Peltier 2013).

Besides the obvious spatial downscaling challenge presented by the basin-scale GRACE TWS retrievals, another challenge for the assimilation of GRACE-based TWS is the partitioning of the vertically integrated TWS retrievals into water cycle component variables. Like the horizontal downscaling of AMSR-E SWE retrievals discussed in the previous section, the partitioning of TWS retrievals can be accomplished through assimilation using an appropriate observation operator. In this case, the observation operator aggregates the fine-scale model estimates of soil moisture, groundwater, and snow to basin-scale TWS estimates. This observation operator enables the computation of the observation-minus-forecast residuals (or innovations) in the (basin-scale, TWS) space of the observations. The observation operator is also needed for the computation of the Kalman gain that transforms the innovations back into the space of the fine-scale model

variables. Similarly, the required temporal aggregation of the model output to the monthly scale of the assimilated TWS retrievals is accomplished through the observation operator.

This concept was illustrated by Forman et al. (2012a), who assimilated GRACE TWS retrievals over the Mackenzie River basin located in northwestern Canada (Figure 3) using an updated version of the GEOS-5 LDAS developed by Zaitchik et al. (2008). The assimilation estimates were evaluated against independent SWE and river discharge observations (section 2.3). Results suggest improved SWE estimates, including improved timing of the subsequent ablation and runoff of the snow pack. For example, Figure 4 shows the improvements in SWE estimates resulting from the assimilation of GRACE TWS retrievals. The white bars represent model results without assimilation, whereas the gray bars represent results with assimilation. The labels on the y-axis of each subplot represent sub-basins of the Mackenzie River basin. As shown in Figure 4, the assimilation of GRACE TWS retrievals generally reduced the mean difference and RMSE between the model and the independent CMC SWE estimates (section 2.3). The reductions are greatest in the Liard basin, where the greatest amount of snow accumulation occurs. Here, the mean difference with the CMC estimates is reduced through GRACE data assimilation by 30% (from 13.2 mm to 9.3 mm) and the RMSE is reduced by 18% (from 24 mm to 19.6 mm). Smaller reductions occur in the other sub-basins. The correlation coefficient of the SWE anomalies (not shown) suggests a slight degree of degradation resulting from assimilation, but further analysis shows there is no statistically significant difference at the 5% level. In summary, the assimilation of GRACE TWS information into the Catchment land surface model reduces the mean difference and RMSE in SWE estimates without adversely impacting estimates of interannual variability.

Additional work was conducted to analyze modeled river discharge estimates against ground-based gauging stations. The findings (not shown) suggest that the assimilation of GRACE observations causes little or no change in the mean difference and RMSE of modeled river discharge, but that small, statistically significant improvements in the anomaly correlations were found. Improvements

in the modeled river runoff anomalies are attributed to a redistribution of the water mass from the snow pack during the accumulation phase into the subsurface during the subsequent ablation and runoff phase. This redistribution of water by the assimilation framework effectively retains water within the hydrologic basin for a longer period of time, which results in small but statistically significant improvements in modeled estimates of river discharge.

Investigation of the analysis increments can provide valuable insights into the behavior of the assimilation procedure and track how much and at what time water is being added to or removed from the individual TWS components. The thin, solid line in Figure 5 shows the increments made to the subsurface water component. Averaged over the Mackenzie River basin and the 7-year experiment period, a total of 12.5 mm of water has been added into the subsurface by the assimilation procedure. This is most evident during the spring and summer. The thick, dashed line in Figure 5 shows the increments for SWE. Averaged over time and space, SWE is removed during the accumulation phase with a small amount added back during the ablation and runoff phase for a total SWE increment of −45.1 mm. Acting together, the analysis increments to the subsurface water and SWE serve to reduce mass during snow accumulation and then increase the mass during ablation and runoff. These two phenomena essentially constrain the amplitude of the modeled TWS dynamics to achieve better agreement of the model estimates with the GRACE retrievals.

The results shown in Figures 4 and 5 imply that the assimilation procedure can effectively partition the vertically integrated GRACE TWS retrievals into their snow and subsurface water components. Houborg et al. (2012), Li et al. (2012), Su et al. (2010), and Zaitchik et al. (2008) further investigated the horizontal, vertical, and temporal disaggregation of GRACE TWS retrievals and reached similar conclusions for other basins in North America and Europe in different climate zones. Collectively, the growing body of research suggests that GRACE TWS assimilation can lead to better understanding of the hydrologic cycle in remote regions of the globe where ground-based observation collection is

difficult, if not impossible. This information could ultimately lead to improved freshwater resource management as well as reduced uncertainty in river discharge.

3.3 Microwave Radiative Transfer Models for Radiance Data Assimilation

It is well established for atmospheric data assimilation systems that the assimilation of satellite radiance observations is preferable to the assimilation of geophysical retrievals (Eyre et al. 1993; Joiner and Dee 2000). The former approach incorporates the radiative transfer model into the assimilation system and thereby avoids inconsistencies in the use of ancillary data between the assimilation system and the (pre-processed) geophysical retrievals. For land data assimilation, however, the vast majority of publications assimilate geophysical retrievals (Lahoz and De Lannoy 2013). In this section, we discuss the development of forward radiative transfer models (RTMs) that convert land surface model variables into microwave brightness temperatures. The first example presents such a model for warm-season microwave brightness temperatures (section 3.3.1). The second example introduces a neural network approach to predict microwave brightness temperatures over snow-covered land (section 3.3.2).

3.3.1 Warm-Season, L-Band Radiative Transfer Modeling

Global observations of brightness temperatures (Tb) at L-band (1.4 GHz) are available from the SMOS mission, and similar Tb observations are expected from the planned Soil Moisture Active Passive (SMAP; Entekhabi et al. 2010) mission. In preparation for the assimilation of Tb observations from SMOS and SMAP, De Lannoy et al. (2012b) added a physically-based, warm-season microwave RTM to the GEOS-5 Catchment model. The RTM is based on the commonly used, zero-order "tau-omega" approach that accounts for microwave emission by the soil and the vegetation canopy as well as attenuation by the vegetation. While the RTM is based on sound physical principles, determining the required parameter values for

493 the microwave roughness, scattering albedo and vegetation optical depth on a
494 global scale is a serious challenge.
495
496 De Lannoy et al. (2012b) collected three different sets of literature values for the
497 L-band RTM parameters. 'Lit1' refers to parameters that are proposed for the
498 future SMAP radiometer retrieval product, 'Lit2' are parameters collected from
499 literature studies using the L-band Microwave Emission of the Biosphere model
500 (Wigneron et al. 2007) and related models, and 'Lit3' is the same as Lit2 except
501 that the microwave roughness parameter is set to values used for SMOS
502 monitoring in the European Centre for Medium-Range Weather Forecasts
503 (ECMWF). The three sets of parameters are illustrated in Figure 6, which shows
504 the resulting microwave roughness (h), vegetation opacity (τ), and scattering
505 albedo (ω) by vegetation class. As can be seen from the figure, there are large
506 differences in h, τ and ω between the three sets of literature values. These
507 differences translate into climatological differences in the simulated brightness
508 temperatures.
509
510 For example, Figure 7a-c shows the differences between one-year mean (1 July
511 2010 to 1 July 2011) model simulations (using the three different literature-based
512 sets of RTM parameters) and SMOS observations for H-polarized Tb at 42.5^{o}
513 incidence angle. Modeled brightness temperatures are at 36 km resolution,
514 commensurate with the resolution of the SMOS observations. Brightness
515 temperatures are screened for frozen soil conditions, snow on the ground, heavy
516 precipitation, proximity to open water surfaces, and radio-frequency interference.
517 The figure shows that all three sets of literature values for the RTM parameters
518 lead to substantial biases against SMOS observations, with Lit1 being too cold (by
519 42.0 K on average) and Lit3 too warm (by 24.6 K on average). Even though Lit2
520 estimates are nearly unbiased in the global average, there are still significant
521 regional biases in the simulated Tbs, with an average absolute bias of 12.7 K.
522 Since such biases would interfere with the assimilation of satellite Tb, the RTM
523 parameters need to be calibrated to achieve climatologically unbiased Tb
524 simulations.
525

526 The most important RTM parameters determining h, τ, and ω have been
527 calibrated, separately for each model grid cell, using multi-angular SMOS
528 observations from 1 July 2011 to 1 July 2012. The calibration simultaneously
529 minimizes, separately for each location, the difference between the modeled and
530 observed climatological mean values, the difference between modeled and
531 observed climatological standard deviations, and the deviations of the optimized
532 parameters from prior guesses (that is, from Lit1, Lit2, or Lit3 values). Through
533 investigating a number of calibration scenarios, De Lannoy et al. (2012b)
534 determined that it is best to simultaneously calibrate a subset of the RTM
535 parameters that most directly determine h, τ and ω.
536
537 After calibration, global Tb simulations for the validation year (1 July 2010 to 1
538 July 2011) are largely unbiased for multiple incidence angles and both H- and V-
539 polarization. For example, Figure 7d shows that the global average absolute bias
540 is now just 2.7 K for H-polarized Tb at 42.5^{o} incidence angle. It should be
541 emphasized that an RMSE of approximately 10 K remains, which is partly due to
542 seasonal biases and partly due to random errors. The former will be addressed in
543 the assimilation system through bias estimation and correction, and the latter
544 through the radiance-based soil moisture analysis. The calibrated parameters are
545 shown in Figure 6. Results suggest, for example, that the roughness parameter (h)
546 is too low in Lit1 and too high in Lit3. The calibrated vegetation opacity (τ) values
547 distinguish clearly between high and low vegetation. The calibrated scattering
548 albedo (ω) is increased over low vegetation, which reduces the vegetation effect in
549 the simulated Tb. In summary, the climatological calibration generates plausible
550 parameter values that are consistent with the underlying land modeling system.
551

552 *3.3.2 Predicting Microwave Brightness Temperatures over Snow*

553 As demonstrated in the previous section, the Catchment model (as do similar
554 global land surface models) supports the application of a physically-based
555 microwave RTM for warm-season processes. However, the snow model
556 components in global land surface models, including that in the Catchment model,
557 are usually too simplistic to support physically-based RTM modeling in the

558 presence of snow. Specifically, global snow models lack reliable estimates of
559 snow microphysical properties (such as grain size, ice layers, and depth hoar)
560 which would be needed for physically-based forward modeling of the microwave
561 brightness temperatures. Forman et al. (2012b) therefore constructed an empirical
562 forward RTM for snow-covered land surfaces based on an Artificial Neural
563 Network (ANN).
564
565 The Catchment model state variables used as input to the ANN include the density
566 and temperature of the snowpack at multiple depths, the temperature of the
567 underlying soil, the overlying air, and the vegetative canopy, and the total amount
568 of water equivalent within the snowpack. In addition, a cumulative temperature
569 gradient index (TGI) is used as a proxy for snow grain size evolution in the
570 presence of a vapor pressure gradient. Using the above inputs, the ANN is trained
571 and (independently) validated using 10.7, 18.7, and 36.5 GHz microwave
572 brightness temperatures at H- and V-polarization from AMSR-E. The
573 independent validation is accomplished as follows: From the 9-year AMSR-E data
574 record, each single year is withheld in turn from the ANN training, and skill
575 metrics for the resulting ANN predictions are computed only against the AMSR-E
576 data that have been withheld from the ANN training.
577
578 Figure 8 demonstrates the performance of the ANN predictions relative to AMSR-
579 E measurements that were not used during training. The figure illustrates the
580 overall ability of the ANN to predict Tbs for the 10 GHz V-polarized channel.
581 The ANN predictions are essentially unbiased (relative to the AMSR-E
582 measurements) across the 9-year period (Figure 8a). The RMSE is typically less
583 than 5 K (Figure 8b). In addition, the ANN demonstrates skill in predicting
584 interannual variability, with anomaly R values well above 0.5 over large parts of
585 North America (Figure 8c). Relatively low skill can be seen in areas along the
586 southern periphery, where the snowpack is relatively thin and ephemeral, as well
587 as in areas north of the boreal forest, where sub-grid scale lake ice (which is not
588 modeled in the land surface model) is common. In short, Figure 8 suggests
589 considerable skill by the ANN at predicting interannual variability in 10 GHz V-
590 polarized Tbs across North America with negligible bias and a reasonable RMSE.

591 The RMSE is somewhat higher but still reasonable (less than 10 K) for the higher
592 frequencies and for H-polarization Tb (see Figures 4-6 of Forman et al. 2012b).
593
594 Forman et al. (2012b) also assessed the potential for using the ANN as a forward
595 observation operator in radiance-based snow assimilation. For this demonstration,
596 the observations are considered to be in the form of spectral differences in V-
597 polarization brightness temperatures, $\Delta Tb \equiv Tb_V(18\ GHz) - Tb_V(36\ GHz)$. Since
598 ΔTb typically increases with increasing SWE, this spectral difference is
599 commonly used to estimate SWE in retrieval algorithms (Kelly 2009). For the
600 demonstration of the radiance-based assimilation considered here, observations of
601 ΔTb imply that the resulting Kalman gain is proportional to error correlations
602 between modeled SWE and ANN predictions of ΔTb. To obtain analysis
603 increments, the Kalman gain would be multiplied with innovations in ΔTb (that is,
604 the difference between actual AMSR-E observations of ΔTb and ANN predictions
605 of ΔTb).
606
607 The Kalman gain computed for 6 February 2003 ranges from -10 mm K^{-1} to 15
608 mm K^{-1} as illustrated in Figure 9. A gain of 1 mm K^{-1} equates to an increase of 1
609 mm in the posterior (updated) modeled SWE for a 1 K innovation (that is, for a
610 difference of 1 K between AMSR-E ΔTb measurements and ANN ΔTb
611 predictions). Similarly, a negative Kalman gain in the presence of a positive-
612 valued innovation would equate to a reduction in modeled SWE. Most
613 importantly, the results suggest that there is a non-zero error correlation between
614 the model SWE forecasts and the simulated ΔTb measurements across much of
615 the North American domain. Overall, the results suggest that the ANN could
616 serve as a computationally efficient observation operator for radiance-based snow
617 data assimilation at the continental scale.
618

619 ### 3.4 Observation Selection for a Root Zone Soil Moisture Analysis

620 Knowledge of the amount of moisture stored in the root zone of the soil is
621 important for many applications related to the transfer of water, energy and carbon
622 between the land and the atmosphere, including the assessment, monitoring, and

prediction of drought (Seneviratne et al. 2010). At the global scale, soil moisture estimates are usually based on two sources of information: (i) direct observations of surface soil moisture from satellite and (ii) observation-based precipitation forcing driving a numerical model of soil moisture dynamics. However, neither surface soil moisture retrievals nor precipitation observations provide direct measurements of soil moisture in the root zone. The selection of the most relevant types of observations for a root zone soil moisture analysis therefore presents an important conceptual problem.

A priori, it is not obvious whether the estimation of root zone soil moisture would benefit more from the use of precipitation observations (as, for example, in the Global Land Data Assimilation System; Rodell et al. 2003) or from the assimilation of surface soil moisture retrievals (as, for example, illustrated by Reichle et al. 2007). This section provides examples of both approaches. First, a land surface reanalysis that relies on observed precipitation is presented, followed by a root zone soil moisture analysis that is based on the assimilation of surface soil moisture retrievals. Finally, the two sources of soil moisture information are merged and compared directly in a single system, and their relative contributions to the skill of root zone soil moisture estimates are assessed.

3.4.1 Using Precipitation Observations

The Modern-Era Retrospective Analysis for Research and Applications (MERRA) is a state-of-the-art atmospheric reanalysis data product based on GEOS-5 that provides, in addition to atmospheric fields, global estimates of soil moisture, latent heat flux, snow, and runoff for 1979-present with a latency of about one month (Rienecker et al. 2011). A supplemental and improved set of land surface hydrological fields ("MERRA-Land") is generated routinely using an improved version of the land component of the MERRA system (Reichle et al. 2011; Reichle 2012). Specifically, the MERRA-Land estimates benefit from corrections to the MERRA precipitation forcing with the global gauge-based NOAA Climate Prediction Center "Unified" (CPCU) precipitation product and from revised

parameter values in the rainfall interception model, changes that effectively correct for known limitations in the MERRA surface meteorological forcings.

With a few exceptions, the MERRA-Land data appear more accurate than the original MERRA estimates and are thus recommended for those interested in using MERRA output for land surface hydrological studies. As an example, Figure 10 examines the drought conditions experienced across the western United States and along the East Coast. The MERRA and MERRA-Land drought indicator shown in the figure is derived by ranking, separately for each grid cell, the normalized, monthly mean root zone soil moisture anomalies for June, July, and August of 1980 through 2011 and converting the rank into percentile units. For comparison, the drought severity assessed independently by U.S. Drought Monitor is also shown. The figure clearly demonstrates that MERRA-Land data are more consistent with the Drought Monitor than MERRA data.

Reichle et al. (2011) and Reichle (2012) provide a more comprehensive and quantitative analysis of the skill (defined as the correlation coefficient of the anomaly time series with independent observations) in land surface hydrological fields from MERRA, MERRA-Land, and the latest global atmospheric reanalysis produced by ECWMF (ERA-I; Dee et al. 2011). Figure 11 shows that MERRA-Land and ERA-I root zone soil moisture skills (against in situ observations at 85 US stations) are comparable and significantly greater than that of MERRA. Furthermore, the runoff skill (against naturalized stream flow observations from 18 US basins) of MERRA-Land is typically higher than that of MERRA and ERA-I (not shown). Throughout the northern hemisphere, MERRA and MERRA-Land agree reasonably well with in situ snow depth measurements (from 583 stations) and with SWE from an independent analysis (not shown). In summary, through observations-based corrections of the MERRA precipitation forcing, MERRA-Land provides a supplemental and significantly improved land surface reanalysis product.

3.4.2 Assimilating Surface Soil Moisture Retrievals

Satellite retrievals of surface soil moisture are not used in MERRA-Land but would almost certainly have further improved the skill of root zone soil moisture estimates. Draper et al. (2012) illustrate the potential gains from assimilating ASCAT (Bartalis et al. 2007; Wagner et al. 1999) and 10.7 GHz AMSR-E Land Parameter Retrieval Model (LPRM; de Jeu et al. 2008) surface soil moisture retrievals. The retrievals are assimilated, both separately and jointly, over 3.5 years into the GEOS-5 LDAS, using MERRA forcing and initial conditions. Soil moisture skill is measured as the anomaly time series correlation coefficient (R) with in situ soil moisture observations from the SCAN/SNOTEL network in the US (66 sites) and the Murrumbidgee Soil Moisture Monitoring Network in Australia (19 sites). These 85 sites are surrounded by terrain with low topographic complexity based on data provided with the ASCAT observations. Averaged over these sites, the ASCAT and AMSR-E surface soil moisture retrievals have similar skill (Draper et al. 2012).

Figure 12 shows the estimated R values and their 95% confidence intervals for root zone soil moisture from the assimilation of ASCAT, AMSR-E and both. The results are benchmarked against an open loop (no assimilation) model integration and have been averaged by land cover type (based on MODIS land cover classifications). Averaged across all 85 sites, assimilating ASCAT and/or AMSR-E surface soil moisture retrievals significantly improved the root zone soil moisture skill (at the 5% level). The mean skill was increased from 0.45 for the open loop, to 0.55 for the assimilation of ASCAT, 0.54 for the assimilation of AMSR-E, and 0.56 for the assimilation of both.

Assimilating the ASCAT or AMSR-E retrievals also improved the mean R value over each individual land cover type, in most cases significantly. At the frequencies observed by AMSR-E and ASCAT, dense vegetation limits the accuracy of soil moisture observations, and so the improvements obtained over the mixed cover sites, which have 10-60% trees or wooded vegetation, are very encouraging. For each land cover type, the skill obtained from the assimilation of ASCAT or AMSR-E retrievals was very similar. The combined assimilation of

718 ASCAT and AMSR-E retrievals generally matched or slightly exceeded the mean
719 R values from the single-sensor assimilation experiments.
720
721 Draper et al. (2012) also examined the contribution of the model skill and the
722 observation skill to the skill of the assimilation estimates. The color surface in
723 Figure 13 shows the skill improvements (ΔR) in root zone soil moisture, where
724 ΔR is defined as the skill (R) of the assimilation estimates (from the single-sensor
725 assimilation of ASCAT or AMSR-E retrievals) minus that of the open loop model
726 estimates. The skill improvements are shown as a function of the open loop
727 model skill and the retrieval skill. Specifically, the ordinate measures the skill of
728 the open loop root zone soil moisture estimates, and the abscissa measures the
729 skill of the assimilated (ASCAT or AMSR-E) surface soil moisture retrievals.
730 Where the skill of the assimilated retrievals is no more than 0.2 less than the open
731 loop skill (below the dashed line), the assimilation improves the root zone soil
732 moisture skill. The improvements increase (up to 0.4) as the observation skill
733 increases relative to that of the open loop (towards the bottom right hand corner).
734 (The results are very similar if the ordinate measures surface soil moisture skill;
735 not shown). Figure 13 thus provides a practical demonstration of the minimum
736 skill required for soil moisture observations to be beneficial in a land data
737 assimilation system and confirms the findings obtained by Reichle et al. (2008b)
738 using synthetically generated observations. In summary, the assimilation of active
739 or passive microwave data significantly improves the model root zone soil
740 moisture estimates by a similar amount, even in cases where the assimilated
741 surface soil moisture retrievals are less skillful than the open loop soil moisture
742 estimates.
743

744 *3.4.3 Combining Precipitation Observations and Surface Soil Moisture*
745 *Retrievals*

746 Liu et al. (2011a) used both precipitation observations and surface soil moisture
747 retrievals within the GEOS-5 LDAS and investigated their relative contributions
748 to the skill of root zone soil moisture estimates. Relative to baseline soil moisture
749 estimates from MERRA, their study investigates soil moisture skill derived from

(i) land model forcing corrections based on large-scale, gauge- and satellite-based precipitation observations and (ii) assimilation of surface soil moisture retrievals from AMSR-E. Three precipitation products were used (separately) to correct the MERRA precipitation towards gauge- and satellite-based observations: the NOAA Climate Prediction Center Merged Analysis of Precipitation (CMAP) pentad product ("standard" version), the Global Precipitation Climatology Project (GPCP) version 2.1 pentad product, and the NOAA Climate Prediction Center (CPC) daily unified precipitation analysis over the United States.

Two different surface soil moisture retrieval products were assimilated into the GEOS-5 LDAS: (i) the operational NASA Level-2B AMSR-E "AE-Land" product (version V09) archived at the National Snow and Ice Data Center (NSIDC; Njoku et al. 2003) and (ii) the AMSR-E LPRM product (de Jeu et al. 2008). Soil moisture skill is assessed using in situ observations in the continental United States at the 37 single-profile sites within the SCAN network for which skillful AMSR-E retrievals are available. As in section 3.4.2, skill is assessed in terms of the anomaly time series correlation coefficient R.

Figure 14 shows comparable average skill for surface soil moisture estimates from the two AMSR-E products and from the Catchment model with MERRA precipitation forcing without data assimilation. Consistent with the findings of section 3.4.1, adding information from precipitation observations increases soil moisture skills for surface and root zone soil moisture. Consistent with the results of section 3.4.2, assimilating satellite estimates of surface soil moisture also increases soil moisture skills, again for surface and root zone soil moisture. The salient result is that adding information from both sources (precipitation observations and surface soil moisture retrievals) increases soil moisture skills by almost the sum of the individual skill contributions, which demonstrates that precipitation corrections and assimilation of satellite soil moisture retrievals contribute important and largely independent amounts of information.

Liu et al. (2011a) also repeated their skill analysis against measurements from four USDA Agricultural Research Service ("CalVal") watersheds with high-

783 quality distributed sensor networks that measure surface soil moisture at the scale
784 of land model and satellite estimates (Jackson et al. 2010). As expected, the skill
785 of the satellite, model, and assimilation estimates is higher when it is assessed
786 against the multi-sensor CalVal observations rather than against single-profile
787 SCAN measurements (not shown). The relative skill contributions by
788 precipitation corrections and soil moisture retrieval assimilation, however, remain
789 unchanged (not shown). This corroborates the results shown in Figure 14 which
790 were obtained with a larger network of single-profile sensors.
791
792 Taken together, the results of this section strongly suggest that future land surface
793 reanalysis efforts would benefit from the use of both precipitation observations
794 and satellite retrievals of surface soil moisture because both types of observations
795 contribute significant and largely independent amounts of information to the skill
796 of root zone soil moisture in the analysis. Moreover, both active and passive
797 surface soil moisture retrievals should be assimilated for maximum coverage and
798 accuracy.
799

4 Summary and Discussion

The present study discussed several conceptual challenges in land surface hydrological data assimilation as part of an effort towards improving our understanding of the Earth's hydrological cycle (Bengtsson 2013). The challenges arise from a seeming mismatch between the assimilated observations and the water cycle variables of interest that can be overcome through the careful design of the assimilation system. This was illustrated with examples from recent research findings using the GEOS-5 LDAS.

The first challenge is the use of coarse-scale satellite observations to estimate land surface fields at finer scales of interest. Such horizontal downscaling can be accomplished by using a fine-scale land surface model and by defining an observation operator that maps from the fine-scale model space to the space of the coarse-scale observations (section 3.1). In the presence of larger-scale model error correlations, the assimilation system can also spread observational information to unobserved locations.

The second challenge is the partitioning of satellite observations (such as TWS retrievals) into their component variables. This partitioning can again be accomplished through an observation operator. In the case of TWS assimilation, the observation operator maps from the fine-scale model estimates of soil moisture and snow to basin-scale TWS (section 3.2). The observation operator therefore enables the computation of the observation-minus-forecast residuals (innovations). The observation operator is also needed for the computation of the Kalman gain matrix that transforms the observation space (coarse-scale TWS) innovations into the model space (fine-scale soil moisture and snow) analysis increments.

The third challenge is the development of microwave RTMs for use as observation operators in radiance-based data assimilation. Two examples were given. In the first example, a global microwave RTM for warm-season, L-band brightness temperatures was calibrated successfully using SMOS observations

(section 3.3.1). In the second example, an empirical approach based on an artificial neural network yielded robust model simulations of AMSR-E microwave brightness temperatures over snow covered land at continental scales (section 3.3.2). In both cases, the results are very encouraging and constitute progress toward replacing the commonly used assimilation of geophysical retrievals (such as SWE or surface soil moisture retrievals) with the direct assimilation of satellite radiances. Note that a radiance-based soil moisture analysis can partition the observational (brightness temperature) information into increments of model soil moisture, soil temperature, and vegetation water content (essentially, the model variables that most impact the brightness temperature). In other words, the microwave RTM, acting as the observation operator, takes on a role that is conceptually similar to that of the observation operator used for the partitioning of TWS information into its water cycle components (section 3.2).

The fourth and final challenge addressed in the paper discusses the selection of the types of observations that are most relevant for the analysis of poorly observed variables. For the analysis of one such variable, root zone soil moisture, the use of gauge- and satellite-based precipitation observations along with active and passive surface soil moisture retrievals was investigated (section 3.4). It was shown that the MERRA-Land surface reanalysis provides better estimates of root zone soil moisture than MERRA due to the use of gauge-based precipitation observations in MERRA-Land. Next, the potential skill gained from the assimilation of surface soil moisture retrievals was investigated. It was demonstrated that improved root zone soil moisture estimates can be obtained even where the skill of the assimilated surface soil moisture retrievals is somewhat poorer than that of the model estimates of surface soil moisture. For maximum coverage and accuracy, both active and passive retrievals should be assimilated. Finally, it was shown that the use of precipitation observations and the assimilation of surface soil moisture retrievals contribute significant and largely independent amounts of root zone soil moisture information. Therefore, future reanalyses should use both of these observation types. This finding is consistent with the general expectation that using more observations in a data assimilation system will improve its output.

In some cases (for example, section 3.1 and section 3.2), the appropriate observation operator and assimilation system configuration entail that neighboring grid cells (or land model tiles) are no longer computationally independent in the assimilation system, even if they are independent in the land model (Reichle and Koster 2003). These computational dependencies arise through spatially correlated perturbations fields or spatially distributed analysis update calculations. Such "three-dimensional" land data assimilation systems therefore necessitate greater computational resources than more simplistic, "one-dimensional" assimilation systems where all model grid cells (or tiles) are treated independently. It is assumed here that the purely technical challenge of computational demand can be overcome with sophisticated software engineering and the increasing availability of affordable and massively parallel computing architectures.

5 Conclusions and Outlook

The present paper focused on the seeming mismatch between satellite observations and the water cycle variables of interest, and how a mismatch can be overcome through careful design and application of a land data assimilation system. Responding to the challenge questions of section 1, we find that, if designed properly, a land data assimilation system can enable

(i) the horizontal downscaling of coarse-scale satellite observations,

(ii) the partitioning of vertically integrated satellite measurements such as TWS into their water cycle components,

(iii) the direct assimilation of satellite radiances for soil moisture or snow analyses, and

(iv) the propagation of information from observed fields such as precipitation and surface soil moisture into variables such as root zone soil moisture, that are of great interest but are not directly observed by satellites.

Naturally, many challenges still lie ahead. State-of-the-art land data assimilation algorithms are only now emerging in operational systems. Much of the recent progress has been achieved in so-called "off-line" (land-only) assimilation systems. These advances need to be incorporated into the coupled land-atmosphere systems used in atmospheric data assimilation and numerical weather prediction (NWP). Ground-breaking advances in coupled land-atmosphere data assimilation are being made, for example, at ECMWF (de Rosnay et al. 2012a, 2012b). At the same time, the coupling of the GEOS-5 LDAS to the GEOS-5 atmospheric data assimilation system is underway at the NASA GMAO.

Moreover, much of the progress in land data assimilation has been with systems that assimilate only one type of observation, often surface soil moisture. In future, more emphasis will need to be placed on the assimilation of multiple types of observations within a single assimilation system, including observations of water cycle components such as soil moisture, SWE, snow cover fraction, TWS, and precipitation.

912 Future development should also address the addition or improvement of runoff
913 routing and surface water storage model components in the global land surface
914 models used in NWP. The planned NASA Surface Water and Ocean Topography
915 (SWOT; Durand et al. 2010) mission, for instance, will provide high-resolution
916 observations of surface water elevation. To improve our understanding of the
917 global hydrological cycle, it will be crucial to incorporate these new observations
918 into global land data assimilation systems, building on early studies such as those
919 by Andreadis et al. (2007), Biancamaria et al. (2011), and Durand et al. (2008).
920
921 Finally, the existing global land data assimilation systems will need to consider
922 the modeling of vegetation dynamics and the assimilation of current or planned
923 satellite observations such as the Fraction of Absorbed Photosynthetically Active
924 Radiation (FAPAR), the Leaf Area Index (LAI), or the multi-angular
925 Photochemical Reflectance Index (PRI) (Albergel et al. 2010; Hilker et al. 2012;
926 Kaminski et al. 2012; Knorr et al. 2010; Munoz-Sabater et al. 2008; Stöckli et al.
927 2011). Furthermore, current microwave sensors already provide observations of
928 the freeze-thaw state of the landscape at coarse scales (Kim et al. 2010), and
929 SMAP will provide much higher-resolution observations with continental
930 coverage (Entekhabi et al. 2010). These vegetation and freeze-thaw observations
931 link the hydrological and carbon cycles and should be used in global land data
932 assimilation systems.
933

934 **Acknowledgments** The authors thank the organizers of the ISSI Workshop on "The Earth's
935 Hydrological Cycle" held February 6-10, 2012 and two anonymous reviewers for their efforts.
936 The research was supported by the NASA program on The Science of Terra and Aqua, the NASA
937 Soil Moisture Active Passive mission, the NASA Postdoctoral Program, and the NASA High-End
938 Computing program.
939
940

References

Albergel C, Calvet J-C, Mahfouf J-F, Rüdiger C, Barbu AL, Lafont S, Roujean J-L, Walker JP, Crapeau M, Wigneron J-P (2010) Monitoring of water and carbon fluxes using a land data assimilation system: a case study for southwestern France. Hydrology and Earth System Sciences 14:1109-1124. doi:10.5194/hess-14-1109-2010

Andreadis K, Lettenmaier D (2006) Assimilating remotely sensed snow observations into a macroscale hydrology model. Advances in Water Resources 29:872–886.

Andreadis KM, Clark EA, Lettenmaier DP, Alsdorf DE (2007) Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. Geophysical Research Letters 34:L10403. doi:10.1029/2007GL029721

Barnett TP, Adam JC, Lettenmaier DP (2005) Potential impacts of a warming climate on water availability in snow-dominated regions. Nature 438:303–309. doi:10.1038/nature04141

Bartalis Z, Wagner W, Naeimi V, Hasenauer S, Scipal K, Bonekamp H, Figa J, Anderson C (2007) Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). Geophysical Research Letters 34:L20401. doi:10.1029/2007GL031088.

Bengtsson L (2013) The response of the global water cycle to climate warming. Surv Geophys, this issue

Biancamaria S, Durand M, Andreadis KM, Bates PD, Boone A, Mognard NM, Rodríguez E, Alsdorf DE, Lettenmaier DP, Clark EA (2011) Assimilation of virtual wide swath altimetry to improve Arctic river modeling. Remote Sensing of Environment 115:373–381. http://dx.doi.org/10.1016/j.rse.2010.09.008

Brasnett B (1999) A global analysis of snow depth for numerical weather prediction. J Appl Meteorol 38:726–740.

Brown RD, Brasnett B (2010) Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data. Environment Canada, Boulder, Colorado, National Snow and Ice Data Center, available at http://nsidc.org/data/docs/daac/nsidc0447_CMC_snow_depth.

Bruinsma S, Lemoine J-M, Biancale R, Vales N (2010) CNES/GRGS 10-day gravity field models (release 2) and the evaluation. Adv Space Res 45:587–601.

Clifford D (2010) Global estimates of snow water equivalent from passive microwave instruments: History, challenges, and future developments. Int J Remote Sens 31:3707–3726.

Crow WT, Wood EF (2003) The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. Adv Water Resour 26:137-149.

Crow WT, Reichle RH (2008) Comparison of adaptive filtering techniques for land surface data assimilation. Water Resources Research 44:W08423. doi:10.1029/2008WR006883

Crow WT, van den Berg MJ (2010) An improved approach for estimating observation and model error parameters for soil moisture data assimilation. Water Resources Research 46:W12519. 10.1029/2010WR009402

979 Dee DP et al. (2011) The ERA-Interim reanalysis: configuration and performance of the data
980 assimilation system. Q J R Meteorol Soc 137:553–597. doi: 10.1002/qj.828
981 de Jeu RAM, Wagner W, Holmes TRH, Dolman AJ, Giesen NC, Friesen J (2008) Global soil
982 moisture patterns observed by space borne microwave radiometers and scatterometers. Surv
983 Geophys 29:399–420. doi:10.1007/s10712-008-9044-0
984 De Lannoy GJM, Reichle RH, Houser PR, PauwelsVRN, Verhoest NEC (2007) Correcting for
985 Forecast Bias in Soil Moisture Assimilation with the Ensemble Kalman Filter. Water Resources
986 Research 43:W09410. doi:10.1029/2006WR005449
987 De Lannoy GJM, Reichle RH, Houser PR, Arsenault KR, Verhoest NEC, Pauwels VRN (2010)
988 Satellite-Scale Snow Water Equivalent Assimilation into a High-Resolution Land Surface
989 Model. Journal of Hydrometeorology 11:352-369. doi:10.1175/2009JHM1192.1
990 De Lannoy GJM, Reichle RH, Arsenault KR, Houser PR, Kumar SV, Verhoest NEC, Pauwels
991 VRN (2012a) Multi-Scale Assimilation of AMSR-E Snow Water Equivalent and MODIS Snow
992 Cover Fraction in Northern Colorado. Water Resources Research, 48, W01522.
993 doi:10.1029/2011WR010588
994 De Lannoy GJM, Reichle RH, Pauwels VRN (2012b) Global Calibration of the GEOS-5 L-band
995 Microwave Radiative Transfer Model over Land Using SMOS Observations. Journal of
996 Hydrometeorology, submitted.
997 de Rosnay P, Drusch M, Vasiljevic D, Balsamo G, Albergel C, Isaksen L (2012a) A simplified
998 Extended Kalman Filter for the global operational soil moisture analysis at ECMWF, Q J R
999 Meteorol Soc, in press. doi:10.1002/qj.2023
1000 de Rosnay P, Balsamo G, Albergel C, Munoz Sabater J, Isaksen L (2012b) Initialisation of land
1001 surface variables for Numerical Weather Prediction. Surv Geophys, in press.
1002 doi:10.1007/s10712-012-9207-x
1003 Draper C S, Reichle RH, De Lannoy GJM, Liu Q (2012) Assimilation of passive and active
1004 microwave soil moisture retrievals. Geophysical Research Letters 39:L04401.
1005 doi:10.1029/2011GL050655
1006 Drusch M (2007) Initializing numerical weather prediction models with satellite derived surface
1007 soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and
1008 the TMI soil moisture data set. J Geophys Res 112: D03102. doi:10.1029/2006JD007478
1009 Drusch M, Wood EF, Gao H (2005) Observation operators for the direct assimilation of TRMM
1010 microwave imager retrieved soil moisture. Geophys Res Let 32:L15403.
1011 doi:10.1029/2005GL023623
1012 Ducharne A, Koster RD, Suarez MJ, Stieglitz M, Kumar P (2000) A catchment-based approach to
1013 modeling land surface processes in a general circulation model, 2: Parameter estimation and
1014 model demonstration. J Geophys Res 105(20):24823-24838.
1015 Dunne S, Entekhabi D (2006) Land surface state and flux estimation using the ensemble Kalman
1016 smoother during the Southern Great Plains 1997 field experiment. Water Resour Res
1017 42:W01407. doi:10.1029/2005WR004334

1018 Durand M, Margulis S (2008) Effects of uncertainty magnitude and accuracy on assimilation of
1019 multi-scale measurements for snowpack characterization. J Geophys Res 113:D02105.
1020 doi:10.1029/2007JD008662
1021 Durand M, Andreadis KM, Alsdorf DE, Lettenmaier DP, Moller D, Wilson MD (2008) Estimation
1022 of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic
1023 model. Geophysical Research Letters 35:L20401. doi:10.1029/2008GL034150
1024 Durand M, Fu LL, Lettenmaier DP, Alsdorf D, Rodríguez E, Esteban-Fernandez D (2010) The
1025 Surface Water and Ocean Topography mission: Observing terrestrial surface water and oceanic
1026 submesoscale eddies. Proceedings of the IEEE 98:766-779.
1027 Ek M, Mitchell K, Yin L, Rogers P, Grunmann P, Koren V, Gayno G, Tarpley JD (2003)
1028 Implementation of Noah land surface model advances in the NCEP operational mesoscale Eta
1029 model. J Geophys Res 108(D22):8851. doi:10.1029/2002JD003296
1030 Entekhabi D et al. (2010) The Soil Moisture Active and Passive (SMAP) Mission. Proceedings of
1031 the IEEE 98:704-716. doi:10.1109/JPROC.2010.2043918
1032 Evensen G (2003) The Ensemble Kalman Filter: theoretical formulation and practical
1033 implementation. Ocean Dynamics 53:343-367. doi:10.1007/s10236-003-0036-9
1034 Eyre JR, Kelly GA, McNally AP, Andersson E, Persson A (1993) Assimilation of TOVS radiance
1035 information through one-dimensional variational analysis. Q J R Meteorol Soc 119:1427-1463.
1036 doi:10.1002/qj.49711951411
1037 Forman BA, Reichle RH, Rodell M (2012a) Assimilation of Terrestrial Water Storage from
1038 GRACE in a Snow-Dominated Basin. Water Resources Research 48:W01507.
1039 doi:10.1029/2011WR011239
1040 Forman BA, Reichle RH, Derksen C (2012b) Estimating Passive Microwave Brightness
1041 Temperature over Snow-covered Land in North America Using a Land Surface Model and an
1042 Artificial Neural Network. IEEE Transactions on Geosciences and Remote Sensing, submitted.
1043 Foster JL, Sun C, Walker JP, Kelly R, Chang A, Dong J, Powell H (2005) Quantifying the
1044 uncertainty in passive microwave snow water equivalent observations. Remote Sens Environ
1045 92(2):187–203.
1046 Foster JL et al. (2011) A blended global snow product using visible, passive microwave and
1047 scatterometer satellite data. Int J Remote Sens 32(5):1371–1395.
1048 doi:10.1080/01431160903548013
1049 Gao Y, Xie H, Lu N, Yao T, Liang T (2010) Toward advanced daily cloud-free snow cover and
1050 snow water equivalent products from Terra-Aqua MODIS and Aqua AMSR-E measurements. J
1051 Hydrol 385:23–35.
1052 Hall DK, Riggs GA (2007) Accuracy assessment of the MODIS snow products. Hydrol Process
1053 21:1534-1547.
1054 Hall DK, Riggs GA, Foster JL, Kumar SV (2010) Development and evaluation of a cloud-gap-
1055 filled MODIS daily snow-cover product. Remote Sens Environ 114:496–503.
1056 Hilker T, Hall FG, Tucker CJ, Coops NC, Black TA, Nichol CJ, Sellers PJ, Barr A, Hollinger DY,
1057 Munger JW (2012) Data assimilation of photosynthetic light-use efficiency using multi-angular

satellite data: II. Model implementation and validation. Remote Sensing of Environment 121:287-300.

Horwath M, Lemoine J-M, Biancale R, Bourgogne S (2011) Improved GRACE science results after adjustment of geometric biases in the Level-1B K-band ranging data. J Geod 85(1):23–38.

Houborg R, Rodell M, Li B, Reichle RH, Zaitchik BF (2012) Drought Indicators Based on Model Assimilated GRACE Terrestrial Water Storage Observations. Water Resources Research 48:W07525. doi:10.1029/2011WR011291

Jackson TJ et al. (2010) Validation of Advanced Microwave Scanning Radiometer Soil Moisture Products. IEEE Trans Geosci Remote Sens 48:4256–4272. doi:10.1109/TGRS.2010.2051035

Joiner J, Dee DP (2000) An error analysis of radiance and suboptimal retrieval assimilation. Q J R Meteorol Soc, 126:1495-1514. doi:10.1002/qj.49712656514

Kaminski T, Knorr W, Scholze M, Gobron N, Pinty B, Giering R, Mathieu P-P (2012) Consistent assimilation of MERIS FAPAR and atmospheric CO2 into a terrestrial vegetation model and interactive mission benefit analysis. Biogeosciences 9:3173-3184. doi:10.5194/bg-9-3173-2012

Kelly RE (2009) The AMSR-E snow depth algorithm: Description and initial results. J Remote Sens Soc Jpn 29:307–317.

Kerr Y, et al. (2010) The SMOS mission: New tool for monitoring key elements of the global water cycle. Proceedings IEEE 98(5):666-687.

Kim Y, Kimball JS, McDonald KC, Glassy J (2010) Developing a Global Data Record of Daily Landscape Freeze/Thaw Status Using Satellite Passive Microwave Remote Sensing. IEEE Trans Geosci Remote Sens 49:949-960. doi:10.1109/TGRS.2010.2070515

Knorr W, Kaminski T, Scholze M, Gobron N, Pinty B, Giering R, Mathieu P-P (2010) Carbon cycle data assimilation with a generic phenology model. J Geophys Res 115:G04017. doi:10.1029/2009JG001119

Knowles KW, Savoie MH, Armstrong RL, Brodzik MJ (2006) updated 2012. AMSR-E/Aqua daily EASE-grid brightness temperatures, September 2002 through May 2011. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.

Koster RD, Suarez MJ, Ducharne A, Stieglitz M, Kumar P (2000) A catchment-based approach to modeling land surface processes in a general circulation model, 1: Model structure. J Geophys Res 105(20):24809-24822.

Kumar SV, Reichle RH, Peters-Lidard CD, Koster RD, Zhan X, Crow WT, Eylander JB, Houser PR (2008a) A Land Surface Data Assimilation Framework using the Land Information System: Description and Applications. Advances in Water Resources 31:1419-1432. doi:10.1016/j.advwatres.2008.01.013

Kumar SV, Peters-Lidard C, Tian Y, Reichle R, Geiger J, Alonge C, Eylander J, Houser P (2008b) An integrated hydrologic modeling and data assimilation framework. IEEE Computer 41:52-59. doi:10.1109/MC.2008.511

Kumar SV, Reichle RH, Harrison KW, Peters-Lidard CD, Yatheendradas S, Santanello JA (2012) A comparison of methods for a priori bias correction in soil moisture data assimilation. Water Resources Research 48:W03515. doi:10.1029/2010WR010261

1098 Lahoz W, De Lannoy GJM (2013) Closing the gaps in our knowledge of the hydrological cycle of
1099 the Earth 8 System: Conceptual Problems. Surv Geophys, this issue
1100 Li B, Rodell M, Zaitchik BF, Reichle RH, Koster RD, van Dam TM (2012) Assimilation of
1101 GRACE Terrestrial Water Storage into a Land Surface Model: Evaluation and Potential Value
1102 for Drought Monitoring in Western and Central Europe. Journal of Hydrology 446-447:103-115.
1103 doi:10.1016/j.jhydrol.2012.04.035
1104 Li L, Gaiser P, Jackson T, Bindlish R, Du J (2007) WindSat soil moisture algorithm and
1105 validation. IEEE International Geoscience and Remote Sensing Symposium 23-28 July 2007,
1106 1188-1191. doi:10.1109/IGARSS.2007.4423017
1107 Liu Q, Reichle RH, Bindlish R, Cosh MH, Crow WT, de Jeu R, De Lannoy GJM, Huffman GJ,
1108 Jackson TJ (2011a) The contributions of precipitation and soil moisture observations to the skill
1109 of soil moisture estimates in a land data assimilation system. Journal of Hydrometeorology
1110 12:750-765. doi:10.1175/JHM-D-10-05000.1
1111 Liu YY, Parinussa RM, Dorigo WA, de Jeu RAM, Wagner W, van Dijk AIJM, McCabe MF,
1112 Evans JP (2011b) Developing an improved soil moisture dataset by blending passive and active
1113 microwave satellite-based retrievals. Hydrol Earth Syst Sci 15(2):425–436. doi:10.5194/hess-
1114 15-425-2011
1115 Margulis SA, McLaughlin D, Entekhabi D, Dunne S (2002) Land data assimilation and estimation
1116 of soil moisture using measurements from the Southern Great Plains 1997 field experiment.
1117 Water Resour Res 38:1299. doi:10.1029/2001WR001114
1118 Muñoz Sabater J, Rüdiger C, Calvet J-C, Fritz N, Jarlan L, Kerr Y (2008) Joint assimilation of
1119 surface soil moisture and LAI observations into a land surface model. Agricultural and Forest
1120 Meteorology 148:1362–1373.
1121 Njoku EG, Jackson TJ, Lakshmi V, Chan TK, Nghiem SV (2003) Soil moisture retrieval from
1122 AMSR-E. IEEE T Geosci Remote Sens 41:215–229. doi:10.1109/TGRS.2002.808243
1123 Pan M, Wood EF (2006) Data Assimilation for Estimating the Terrestrial Water Budget Using a
1124 Constrained Ensemble Kalman Filter. J Hydrometeorol 7:534-547.
1125 Pan M, Wood EF, Wojcik R, McCabe MF (2008) Estimation of regional terrestrial water cycle
1126 using multi-sensor remote sensing observations and data assimilation. Remote Sensing of
1127 Environment 112:1282-1294. doi:10.1016/j.rse.2007.02.039
1128 Parinussa RM, Holmes TRH, de Jeu RAM (2012) Soil moisture retrievals from the WindSat
1129 spaceborne polarimetric microwave radiometer. IEEE Trans Geosci Remote Sens, in press.
1130 doi:10.1109/TGRS.2011.2174643
1131 Peltier WR (2013) Application of GRACE time-dependent gravity measurements to the
1132 understanding of land surface hydrology. Surv Geophys, this issue
1133 Pulliainen J (2006) Mapping of snow water equivalent and snow depth in boreal and subarctic
1134 zones by assimilating space-borne microwave radiometer data and ground-based observations.
1135 Remote Sensing of Environment 101:257–269.

1136 Reichle RH, Entekhabi D, McLaughlin DB (2001) Downscaling of radiobrightness measurements
1137 for soil moisture estimation: A four-dimensional variational data assimilation approach. Water
1138 Resources Research 37:2353-2364. doi:10.1029/2001WR000475
1139 Reichle RH, McLaughlin DB, Entekhabi D (2002a) Hydrologic data assimilation with the
1140 Ensemble Kalman filter. Monthly Weather Review 130:103-114.
1141 Reichle RH, Walker JP, Koster RD, Houser PR (2002b) Extended vs. Ensemble Kalman Filtering
1142 for Land Data Assimilation. Journal of Hydrometeorology 3:728-740.
1143 Reichle RH, RD Koster (2003) Assessing the impact of horizontal error correlations in background
1144 fields on soil moisture estimation. Journal of Hydrometeorology 4:1229-1242.
1145 Reichle RH, RD Koster (2004) Bias reduction in short records of satellite soil moisture.
1146 Geophysical Research Letters 31:L19501. doi:10.1029/2004GL020938
1147 Reichle RH, RD Koster (2005) Global assimilation of satellite surface soil moisture retrievals into
1148 the NASA Catchment land surface model. Geophysical Research Letters 32:L02404.
1149 doi:10.1029/2004GL021700
1150 Reichle RH, Koster RD, Liu P, Mahanama SPP, Njoku EG, Owe M (2007) Comparison and
1151 assimilation of global soil moisture retrievals from the Advanced Microwave Scanning
1152 Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel
1153 Microwave Radiometer (SMMR). Journal of Geophysical Research 112: D09108.
1154 doi:10.1029/2006JD008033
1155 Reichle RH (2008) Data Assimilation Methods in the Earth Sciences. Advances in Water
1156 Resources 31:1411-1418. doi:10.1016/j.advwatres.2008.01.001
1157 Reichle RH, Crow WT, Koster RD, Sharif H, Mahanama SPP (2008a) Contribution of soil
1158 moisture retrievals to land data assimilation products. Geophysical Research Letters 35:L01404.
1159 doi:10.1029/2007GL031986
1160 Reichle RH, Crow WT, Keppenne CL (2008b) An adaptive ensemble Kalman filter for soil
1161 moisture data assimilation. Water Resources Research 44:W03423.
1162 doi:10.1029/2007WR006357
1163 Reichle RH, Bosilovich MG, Crow WT, Koster RD, Kumar SV, Mahanama SPP, Zaitchik BF
1164 (2009) Recent Advances in Land Data Assimilation at the NASA Global Modeling and
1165 Assimilation Office. In: Data Assimilation for Atmospheric, Oceanic and Hydrologic
1166 Applications, edited by Seon K Park and Liang Xu, 407-428, Springer Verlag, New York.
1167 doi:10.1007/978-3-540-71056-1
1168 Reichle RH, Koster RD, De Lannoy GJM, Forman BA, Liu Q, Mahanama SPP, Toure A (2011)
1169 Assessment and enhancement of MERRA land surface hydrology estimates. Journal of Climate
1170 24:6322-6338. doi:10.1175/JCLI-D-10-05033.1
1171 Reichle RH (2012) The MERRA-Land Data Product (Version 1.1). NASA Global Modeling and
1172 Assimilation Office, Office Note No. 3. Available at http://gmao.gsfc.nasa.gov/pubs/
1173 Rienecker MM et al. (2008) The GEOS-5 Data Assimilation System—Documentation of versions
1174 5.0.1 and 5.1.0, and 5.2.0. NASA Tech. Rep. Series on Global Modeling and Data Assimilation,
1175 NASA/TM-2008-104606, Vol 27, 92 pp. Available at http://gmao.gsfc.nasa.gov/pubs/

1176 Rienecker MM et al. (2011) MERRA - NASA's Modern-Era Retrospective Analysis for Research
1177 and Applications. Journal of Climate 24:3624-3648. doi:10.1175/JCLI-D-11-00015.1
1178 Rodell M et al. (2003) The global land data assimilation system. Bull Am Meteorol Soc 85:381-
1179 394. doi:10.1175/BAMS-85-3-381
1180 Rodell M, Velicogna I, Famiglietti JS (2009) Satellite-based estimates of groundwater depletion in
1181 India. Nature 460:999-1002. doi:10.1038/nature08238
1182 Rott H (2013) The Role of Snow in the Water Cycle and Climate System – Observations and
1183 Modelling. Surv Geophys, this issue
1184 Rowlands DD, Luthcke SB, Klosko SM, Lemoine FGR, Chinn DS, McCarthy JJ, Cox CM,
1185 Anderson OB (2005) Resolving mass flux at high spatial and temporal resolution using GRACE
1186 intersatellite measurements. Geophys Res Lett 32:L04310. doi:10.1029/2004GL021908
1187 Rowlands DD, Luthcke SB, McCarthy JJ, Klosko SM, Chinn DS, Lemoine FG, Boy J-P, Sabaka
1188 TJ (2010) Global mass flux solutions from GRACE: A comparison of parameter estimation
1189 strategies—Mass concentrations versus Stokes coefficients. J Geophys Res 115:B01403.
1190 doi:10.1029/2009JB006546
1191 Sahoo AK, De Lannoy GJM, Reichle RH, Houser PR (2012) Assimilation and Downscaling of
1192 Satellite Observed Soil Moisture over the Little River Experimental Watershed in Georgia,
1193 USA. Advances in Water Resources, in press.
1194 Schaefer GL, Cosh MH, Jackson TJ (2007) The USDA Natural Resources Conservation Service
1195 Soil Climate Analysis Network (SCAN). J Atmos Oceanic Technol 24:2073–2077.
1196 doi:10.1175/2007JTECHA930.1
1197 Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner I, Orlowsky B, Teuling AJ
1198 (2010) Investigating soil moisture–climate interactions in a changing climate: A review. Earth-
1199 Science Reviews 99:125–161. doi:10.1016/j.earscirev.2010.02.004
1200 Smith AB et al. (2012) The Murrumbidgee soil moisture monitoring network data set. Water
1201 Resour Res 48:W07701. doi:10.1029/2012WR011976
1202 Stieglitz M, Ducharne A, Koster RD, Suarez M (2001) The impact of detailed snow physics on the
1203 simulation of snow cover and subsurface thermodynamics at continental scales. J
1204 Hydrometeorol 2:228–242.
1205 Stöckli R, Rutishauser T, Baker I, Liniger MA, Denning AS (2011) A global reanalysis of
1206 vegetation phenology. J Geophys Res 116:G03020. doi:10.1029/2010JG001545
1207 Sturm M, Holmgren J, Liston GE (1995) A Seasonal Snow Cover Classification System for Local
1208 to Global Applications. J Climate 8:1261-1283.
1209 Sturm M, Taras B, Liston GE, Derksen C, Jonas T, Lea J (2010) Estimating snow water equivalent
1210 using snow depth data and climate classes. J Hydrometeorol 11:1380–1394.
1211 Su H, Yang Z-L, Niu G-Y, Dickinson RE (2008) Enhancing the estimation of continental-scale
1212 snow water equivalent by assimilating MODIS snow cover with the ensemble Kalman filter. J
1213 Geophys Res 113:D08120. doi:10.1029/2007JD009232

Su H, Yang Z-L, Dickinson RE, Wilson CR, Niu G-Y (2010) Multisensor snow data assimilation at the continental scale: The value of gravity recovery and climate experiment terrestrial water storage information. J Geophys Res 115:D10104. doi:10.1029/2009JD013035

Swenson S, Wahr J (2006) Post-processing removal of correlated errors in GRACE data. Geophys Res Lett 33:L08402. doi:10.1029/2005GL025285

Tang Q, Gao H, Yeh P, Oki T, Su F, Lettenmaier DP (2010) Dynamics of terrestrial water storage change from satellite and surface observations and modeling. J Hydrometeorol 11:156-170.

Tedesco M, Narvekar PS (2010) Assessment of the NASA AMSR-E SWE product. IEEE J Sel Top Appl Earth Obs Remote Sens 3:141–159.

Tedesco M, Reichle R, Löw A, Markus T, Foster JL (2010) Dynamic approaches for snow depth retrieval from spaceborne microwave brightness temperature. IEEE Trans Geosci Remote Sens, 48(4):1955–1967.

Trenberth K (2013) Challenges for the global water cycle. Surv Geophys, this issue

Wagner W, Lemoine G, Rott H (1999) A method for estimating soil moisture from ERS scatterometer and soil data. Remote Sens Environ 70:191-207. doi:10.1016/S0034-4257(99)00036-X

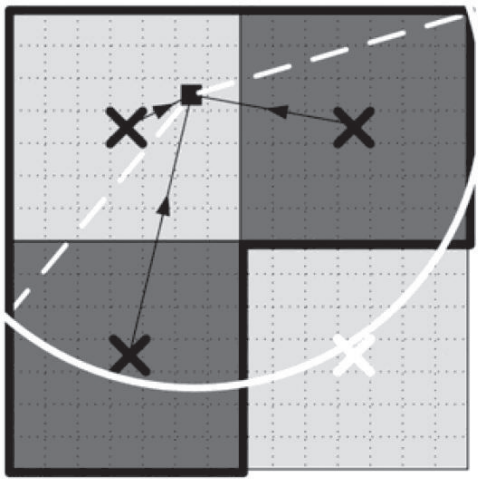
Wahr J, Swenson S, Zlotnicki V, Velicogna I (2004) Time-variable gravity from GRACE: First results. Geophys Res Lett 31:L11501. doi:10.1029/2004GL019779

Wigneron J-P et al. (2007) L-band microwave emission of the biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields. Remote Sensing of Environment 107:639–655.

Zaitchik BF, Rodell M, Reichle RH (2008) Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin. Journal of Hydrometeorology 9:535-548. doi:10.1175/2007JHM951.1

Zhou Y, McLaughlin D, Entekhabi D (2006) Assessing the Performance of the Ensemble Kalman Filter for Land Surface Data Assimilation. Mon Weather Rev 134:2128-2142.

1241 # Figures



1242

1243 **Fig. 1** Schematic of the distributed ("three-dimensional") EnKF update used for the assimilation
1244 of coarse-scale snow observations. See text for details. Adapted from (De Lannoy et al. 2010).
1245

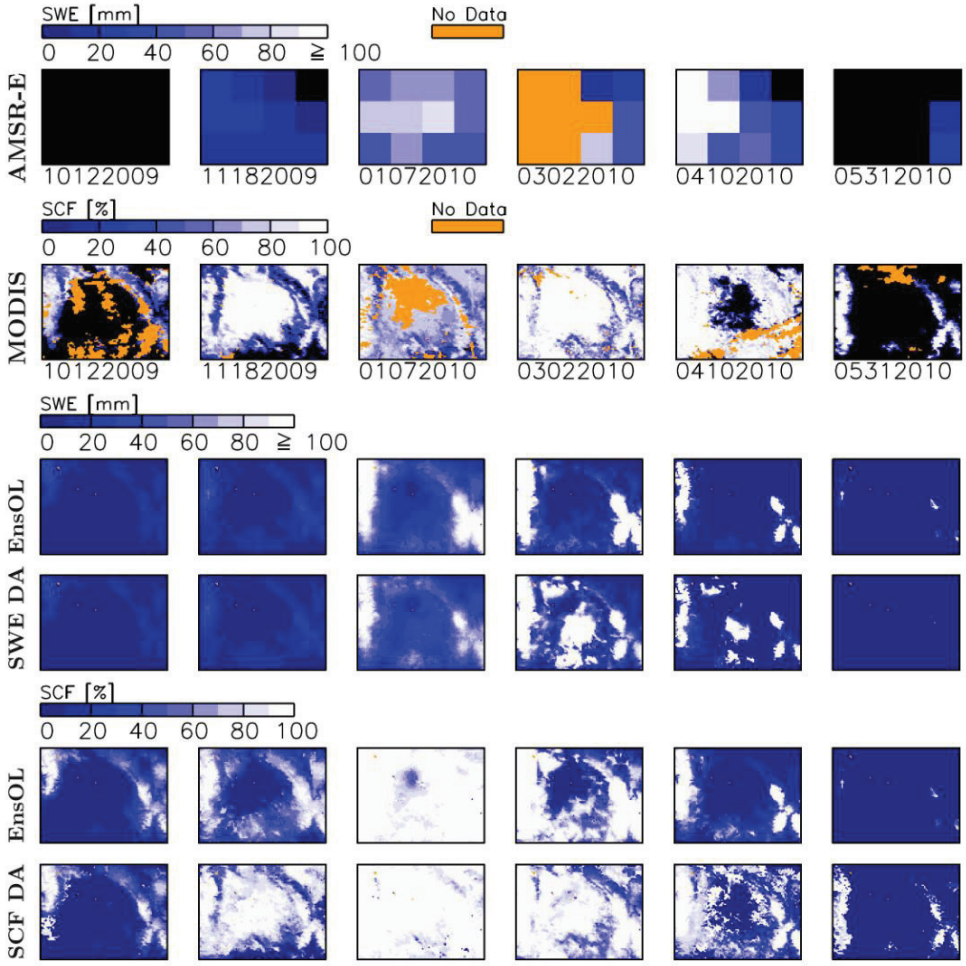
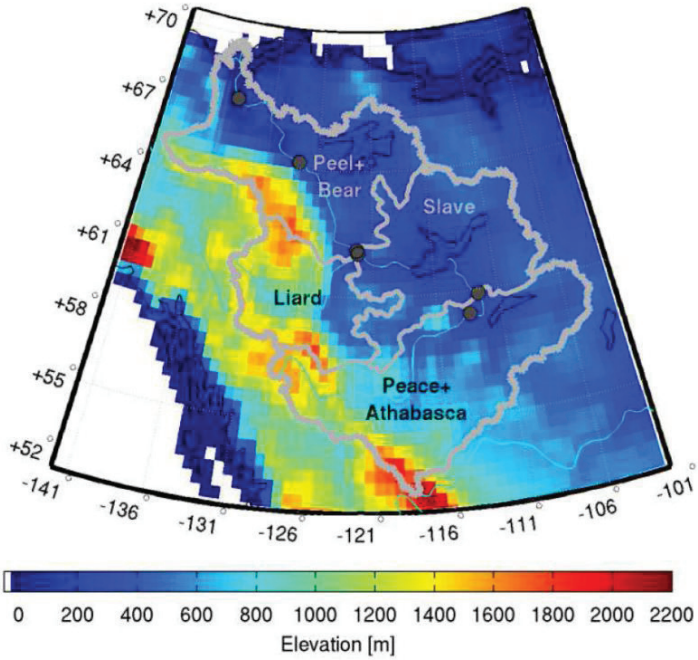


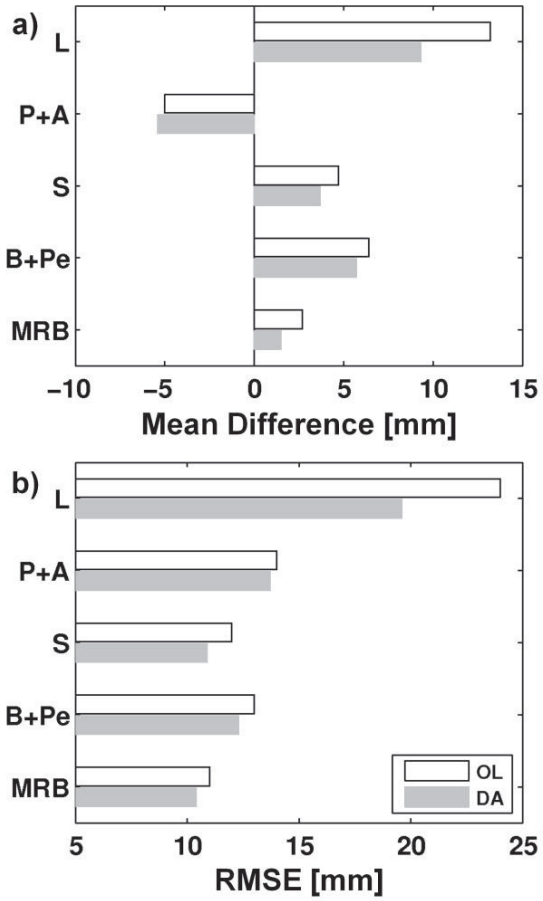
Fig. 2 SWE and SCF fields for 6 days (MMDDYYYY) in the winter of 2009–2010 for a 75 km by 100 km domain (1 km resolution) in northern Colorado. Blue (white) colors indicate low (high) SWE or SCF, black shading indicates no snow, and orange shading indicates no data. The top two rows show SWE and SCF satellite observations. The remaining rows show SWE (rows 3 and 4) and SCF (rows 5 and 6) for the ensemble Open Loop (EnsOL) forecast (no assimilation) and the analyses obtained through data assimilation (DA) of SWE or SCF. Adapted from (De Lannoy et al. 2012a).



1256

1257 **Fig. 3** Map of the 1,800,000 km^2 Mackenzie River Basin including GEOS-5 topography, sub-
1258 basin delineation, and GRDC observation locations (solid dots). Adapted from (Forman et al.
1259 2012a).

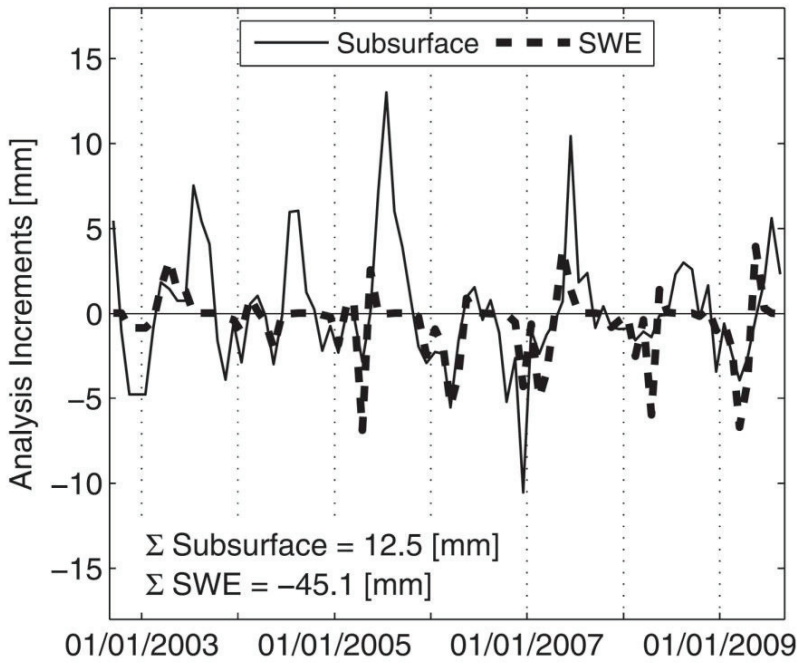
1260



1261

1262 **Fig. 4** SWE statistics of a) mean difference and b) RMSE for open loop (OL; white) and
1263 assimilation (DA; light gray) of GRACE TWS retrievals relative to CMC SWE estimates via
1264 Sturm el al. (2010). Statistics are for the Mackenzie River Basin (MRB) and its subbasins Liard
1265 (L), Peace and Athabasca (P+A), Slave (S), and Bear and Peel (B+Pe) shown in Figure 3.
1266 Adapted from (Forman et al. 2012a).

1267



1268

1269 **Fig. 5** Analysis increments for the entire Mackenzie River basin from GRACE TWS assimilation.
1270 The thin, solid line represents the subsurface water increments whereas the thick, dashed line
1271 represents the SWE increments. From (Forman et al. 2012a).

1272

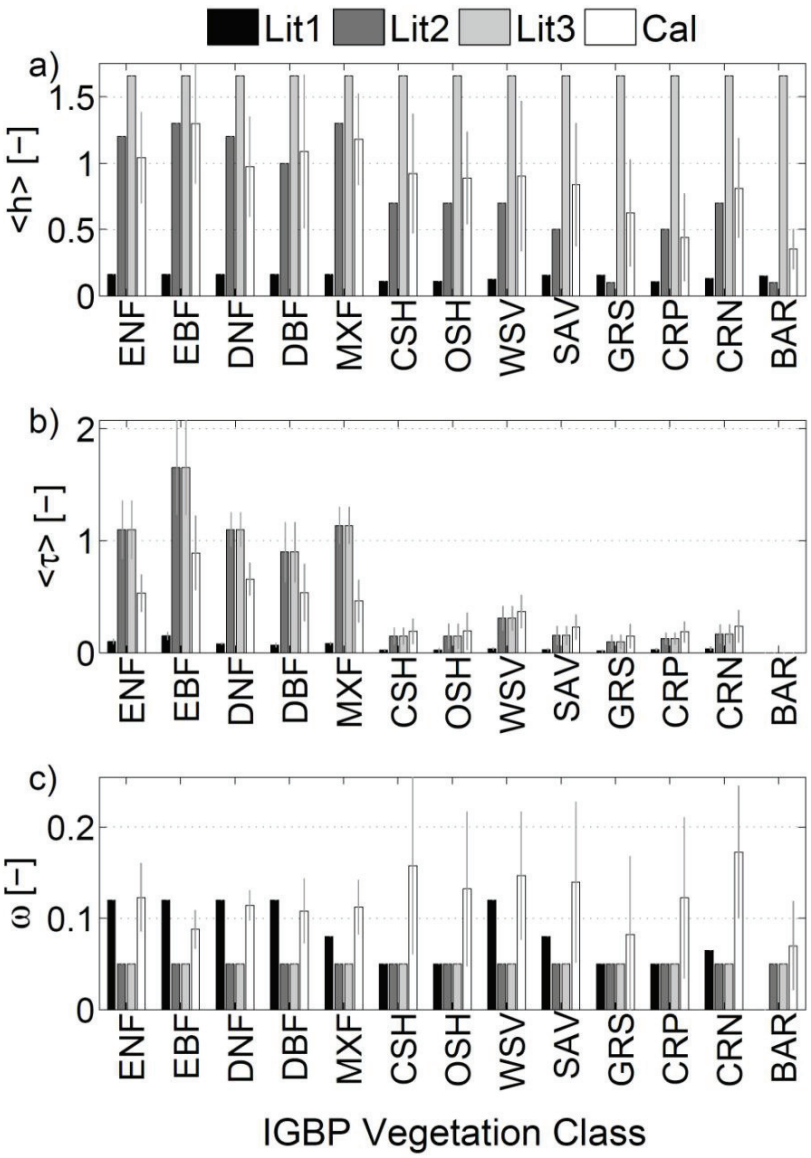
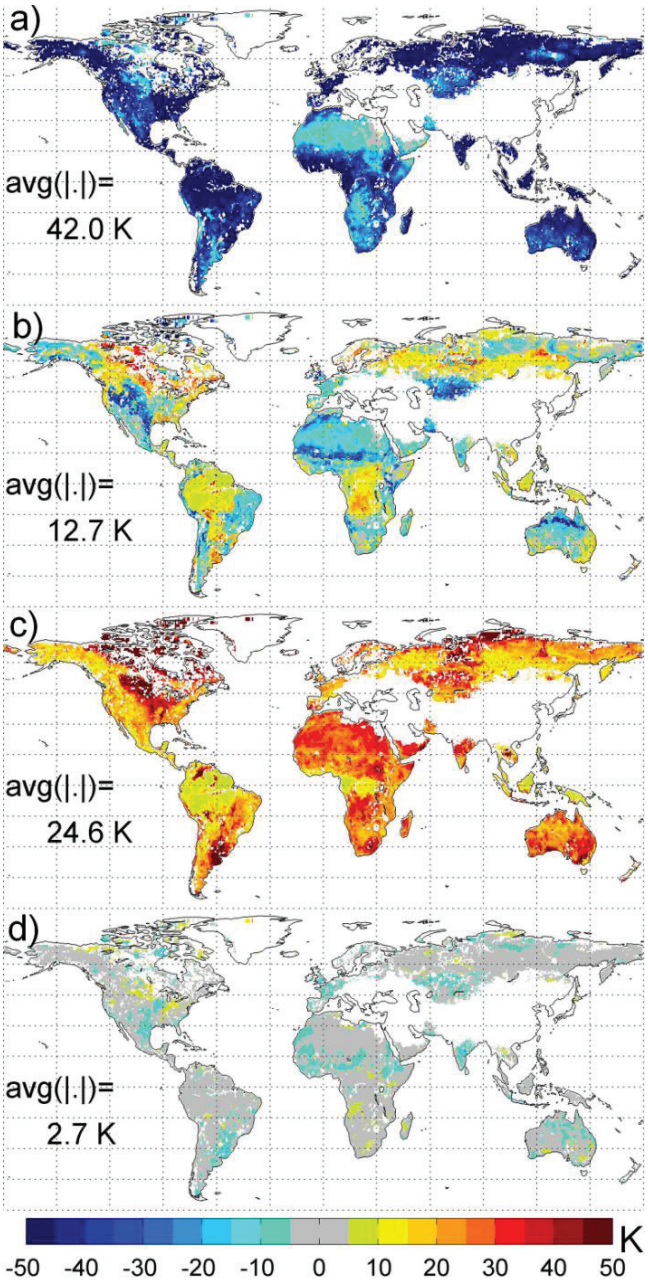


Fig. 6 (a) Time-mean <*h*> (1 July 2010 – 1 July 2011), (b) time-mean <τ>, and (c) time-invariant ω; (Lit1, Lit2 and Lit3) before calibration, and (Cal) after calibration, spatially averaged by vegetation class. International Geosphere-Biosphere Programme (IGBP) vegetation classes are (ENF) Evergreen Needleleaf Forest, (EBF) Evergreen Broadleaf Forest, (DNF) Deciduous Needleleaf Forest, (DBF) Deciduous Broadleaf Forest, (MXF) Mixed Forest, (CSH) Closed Shrublands, (OSH) Open Shrublands, (WSV) Woody Savannas, (SAV) Savannas, (GRS) Grasslands, (CRP) Croplands, (CRN) Cropland and Natural Vegetation, and (BAR) Barren or Sparsely Vegetated. Thin gray lines for Cal indicate the spatial standard deviation within each vegetation class. From (De Lannoy et al. 2012b).



1283

1284 **Fig. 7** Difference between one-year (1 July 2010 – 1 July 2011) mean values of $Tb_H(42.5°)$ in
1285 Kelvin from GEOS-5 and SMOS observations for (a) Lit1, (b) Lit2, (c) Lit3, and (d) calibrated
1286 parameters. Within each subplot, "avg(|.|)" indicates the average absolute difference across the
1287 globe (excluding regions impacted by open water or radio-frequency interference that are shown in
1288 white). Adapted from (De Lannoy et al. 2012b).

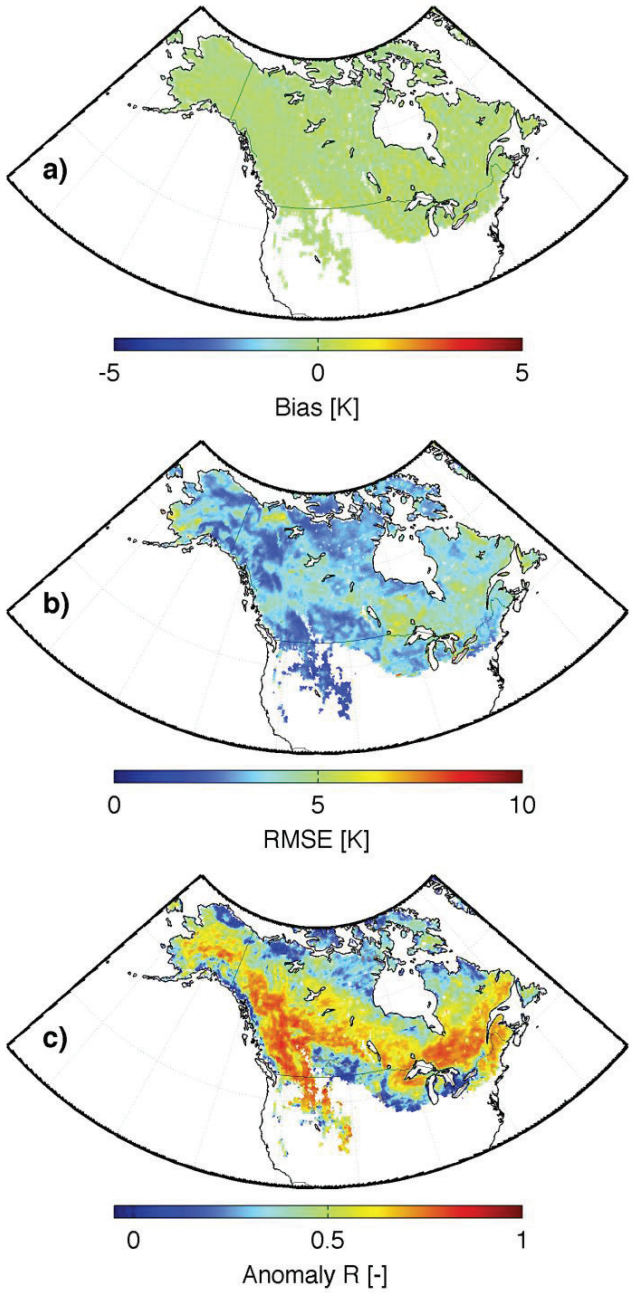


Fig. 8 (a) Bias, (b) RMSE, and (c) anomaly R for ANN simulated 10 GHz V-polarized Tb from 1 September 2002 to 1 September 2011 vs. AMSR-E observations not used in training. Anomaly R values not statistically different from zero at the 95% significance level based on a Fisher Z transform are shown in gray. Such non-significant R values occur in only a few very small regions.

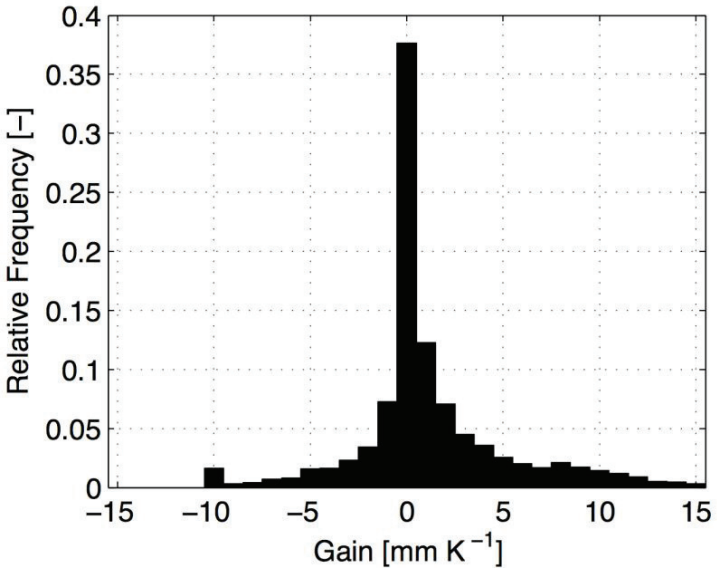
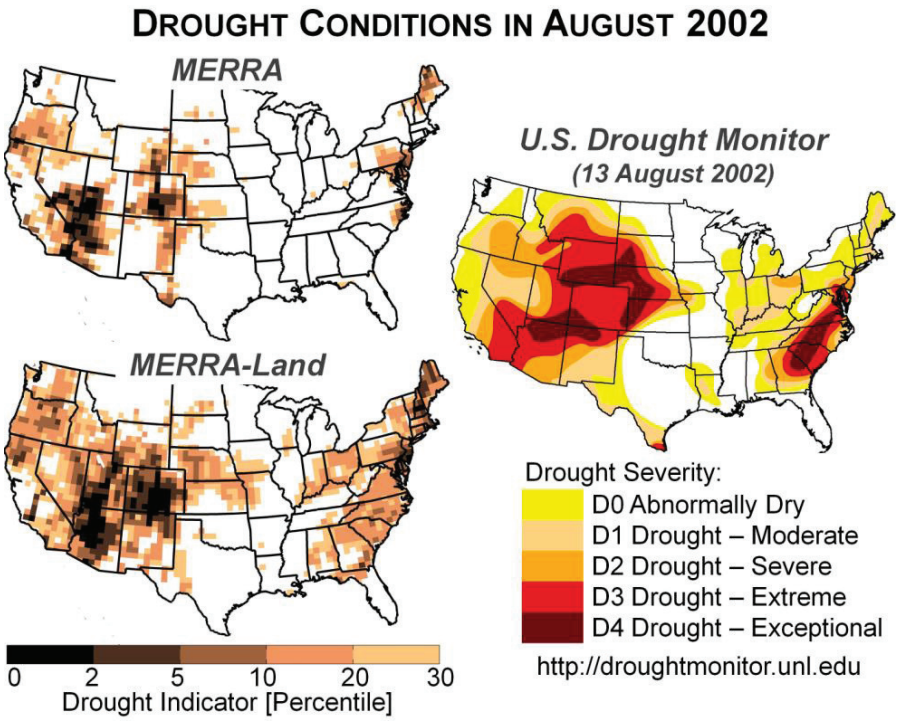
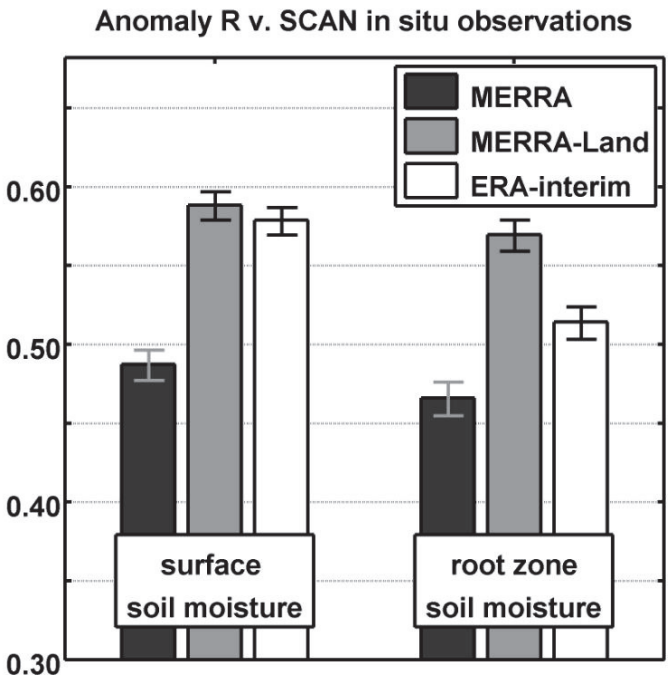


Fig. 9 Histogram of the Kalman gain on 6 February 2003 for SWE versus $\Delta Tb=(Tb_V(18\ GHz)-Tb_V(36\ GHz))$.



1300

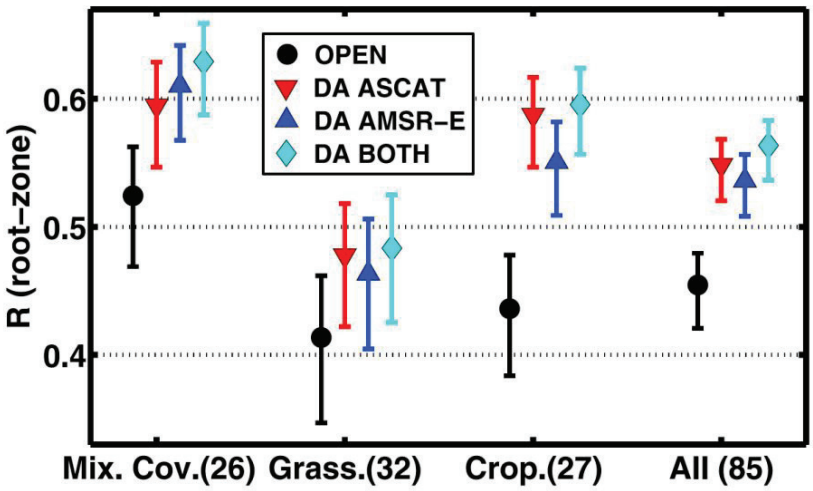
1301 **Fig. 10** Drought indicator derived from (top left) MERRA and (bottom left) MERRA-Land root
1302 zone soil moisture estimates for August 2002. Darker colors indicate more severe drought
1303 conditions. MERRA-Land estimates are more consistent than MERRA estimates with an
1304 independent drought assessment from the US Drought Monitor for 13 August 2002 (right).
1305



1306

1307 **Fig. 11** Skill (pentad anomaly R; dimensionless) of MERRA, MERRA-Land, and ERA-I
1308 estimates (2002-2009) versus SCAN in situ surface and root zone soil moisture measurements at
1309 85 stations. Error bars indicate approximate 95% confidence intervals. Adapted from (Reichle et
1310 al. 2012).

1311



1312

1313 **Fig. 12** Mean skill for root zone soil moisture from the open loop (ensemble mean, no
1314 assimilation), and the data assimilation (DA) of ASCAT, AMSR-E, and both surface soil moisture
1315 retrievals, averaged by land cover class, with 95% confidence intervals. The number of sites in
1316 each land cover class is given in the axis labels. Skill is defined as the daily anomaly R value vs.
1317 SCAN/SNOTEL and Murrumbidgee in situ observations. Adapted from (Draper et al. 2012).
1318

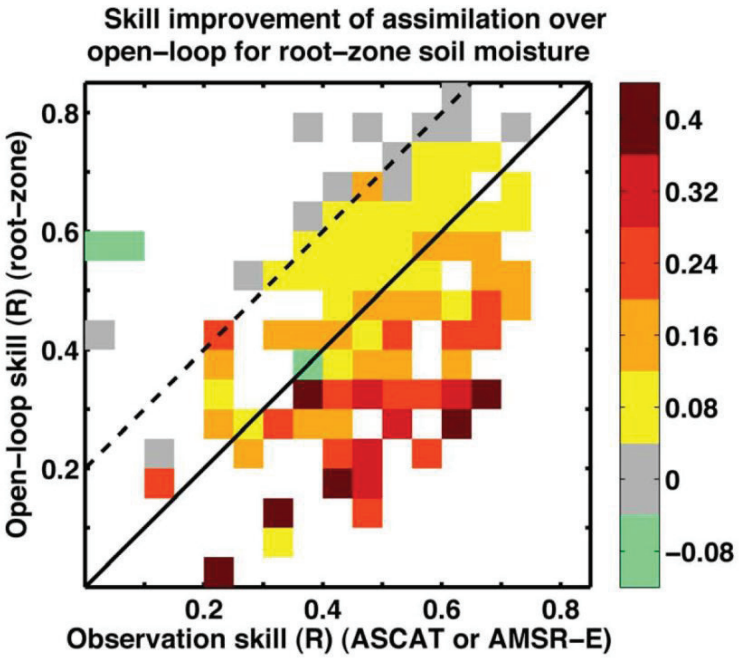
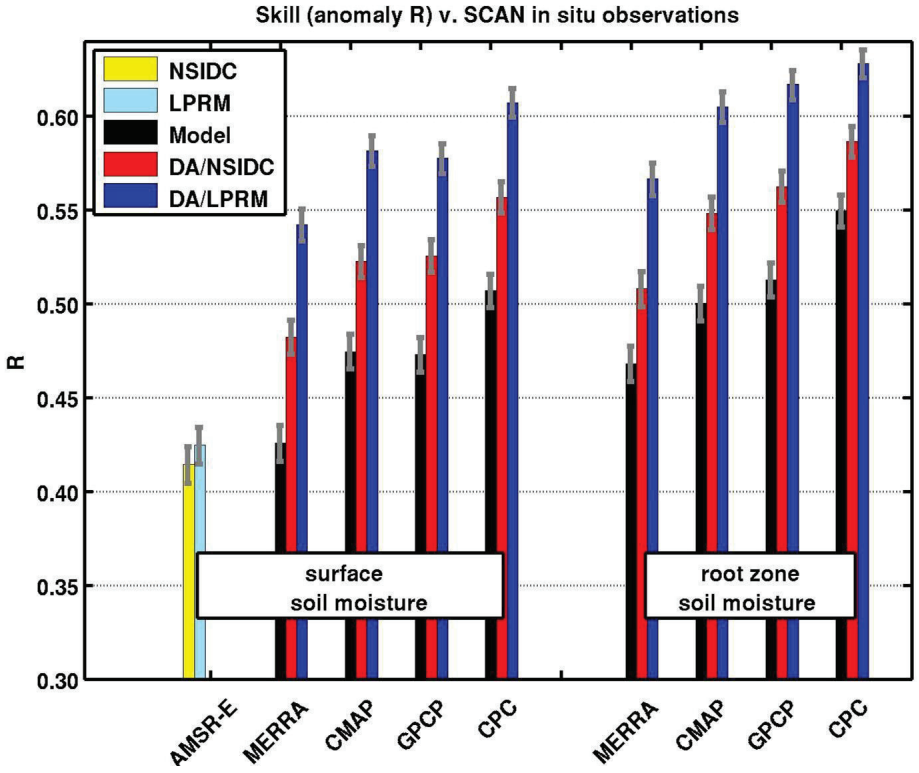


Fig. 13 Root zone soil moisture skill improvement (ΔR) from assimilating either ASCAT or AMSR-E surface soil moisture retrievals as a function of (ordinate) the open loop model skill and (abscissa) the observation skill. Skill improvement (ΔR) is defined as the skill of the assimilation product minus the open loop skill, with skill based only on days with data available from both satellites. Skill is assessed versus in situ measurements from the SCAN and Murrumbidgee networks. Significant improvements are found in the area below the dashed line where the skill of the retrievals may be lower than that of the open loop by up to 0.2. Adapted from (Draper et al. 2012).



1329

1330 **Fig. 14** Skill (daily anomaly R; dimensionless) vs. SCAN in situ soil moisture measurements for
1331 estimates from two AMSR-E retrieval datasets (NSIDC and LPRM), the Catchment model forced
1332 with four different precipitation datasets (MERRA, CMAP, GPCP, and CPC), and the
1333 corresponding data assimilation integrations (red bars: DA/NSIDC and blue bars: DA/LPRM).
1334 Average is based on 37 SCAN sites for surface and 35 SCAN sites for root zone soil moisture.
1335 Error bars indicate approximate 95% confidence. From (Liu et al. 2011a).